

The Sonic Dimension as Dramatic Driver in 21st-Century Pop Music

by

Asaf Peres

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Music Theory)
in the University of Michigan
2016

Doctoral Committee:

Professor Marion A. Guck, Chair
Associate Professor Jason A. Corey
Associate Professor Karen J. Fournier
Professor Patricia Hall
Professor Andrew W. Mead, Indiana University
Associate Professor Derek W. Vaillant

© Asaf Peres

All rights reserved

2016

To my amazing parents, Lily and Mordechai, who inspired me and helped me to follow my dreams from day one. And to my incredible wife, Rebecca, who provided me not only with love and emotional support, but also with an abundance of intellectual support, and whose presence in my life contributed greatly to my ability to complete this dissertation without pulling a single all-nighter.

Acknowledgements

I owe a huge debt of gratitude to Marion Guck, who went above and beyond her duties as my committee chair. Professor Guck was always available for meetings, Skype calls, and email exchanges, and was consistently encouraging and enthusiastic about this project. She provided me with invaluable guidance, generously shared her knowledge and experience, and approached my ideas with an open mind, excitedly reacting to the good ones while challenging me to reexamine those that needed fine-tuning.

Prior to beginning my work on this project, I conducted an independent study under the mentorship of Jason Corey, who helped me to learn the fundamentals of using digital audio workstations, mixing and mastering, reverse engineering, and reading spectrograms. Acquiring those skills was crucial to my ability to conduct this research. Professor Corey also introduced me to much of the literature on music technology and sound production, which informed my early ideas.

In my numerous conversations with Andrew Mead, he always offered encouragement and guidance, and exhibited an astounding ability to provide me with insights into my own ideas. His advice helped me to refine and clarify many of the concepts in this research. Karen Fournier and Derek Vaillant also provided support and unique perspectives that contributed to the evolution of this work.

The University of Michigan's Institute for the Humanities fellowship program both funded this research and gave me the opportunity to dialogue with a wonderful group of scholars from a wide range of disciplines, which helped me to expand my intellectual horizons and communicate my ideas to diverse

audiences. I was additionally supported by the Rackham Graduate School's Predoctoral fellowship, as well as Rackham travel and research grants.

This dissertation would not exist without the songwriters, producers, artists, and sound engineers who created the songs that I analyzed, as well as the developers of the technological tools that I used for analyzing these songs and their sounds – *Ableton Live*, *Sonic Visualizer*, and *Variations Audio Timeliner*.

Last but certainly not least, my wife, Rebecca, who read many of my early drafts, gave me valuable advice, and taught me how to be productive even before the sun goes down. Although I am very proud of this dissertation, marrying Rebecca was my finest achievement as a graduate student.

Table of Contents

| | |
|---|-----|
| Dedication..... | ii |
| Acknowledgements..... | iii |
| List of Examples..... | vii |
| List of Abbreviations..... | xi |
| Abstract..... | xii |
| Chapter 1: Introduction: Sonic Syntax and Analysis of Contemporary Pop Music..... | 1 |
| Chapter Overview..... | 7 |
| Chapter 2: Challenges and Methodology..... | 11 |
| General Challenges Related to Analyzing Contemporary Pop..... | 12 |
| Visual Representations..... | 28 |
| Summary..... | 36 |
| Chapter 3: (Dys)Functional Harmony: Handing Over the Role of Harmonic Function to Sound Production..... | 37 |
| The Evolution of Pop Music..... | 40 |
| The Components of Anticipation..... | 47 |
| Anticipatory Functions..... | 56 |
| Conclusion..... | 72 |
| Chapter 4: The Anticipation Chain: Creating Tension and Release by Manipulating Sonic Density, Sonic Energy, and Rhythmic Intensity..... | 73 |
| Techniques..... | 74 |
| Resolving/Defusing Tension..... | 101 |
| The Big Picture..... | 111 |
| Conclusion..... | 123 |

| | |
|--|-----|
| Chapter 5: Verse-Chorus Song Form, as Defined by Modern Sound Production.. | 126 |
| Verse-Chorus Form..... | 132 |
| Distinguishing Between the Post-Chorus, Bridge, and Transition..... | 147 |
| Exploring More Complex Forms..... | 165 |
| Conclusion..... | 182 |
| Chapter 6: Conclusion..... | 183 |
| Appendix - Analyzed Sample of Pop Songs 2008-2016..... | 187 |
| Bibliography..... | 191 |
| Recordings..... | 196 |

List of Examples

| Example No. | Title | Timing | Type | Page |
|--------------------|--|--|--------------------------|-------------|
| 2.1 | Stages of a typical filter sweep | N/A | DAW screen capture | 18 |
| 2.2 | Envelope settings in <i>Ableton Live's</i> Operator synthesizer | N/A | DAW screen capture | 21 |
| 2.3 | Harmonic partials settings in <i>Ableton Live's</i> Operator synthesizer | N/A | DAW screen capture | 22 |
| 2.4 | Ariana Grande, "Break Free" | N/A | Timeline | 29 |
| 2.5 | Britney Spears, "Till the World Ends" | 2:49-3:23 | Spectrogram | 32 |
| 2.6 | Alesso, "Heroes" | 0:53-1:10 | Spectrogram | 34 |
| 2.7 | Ariana Grande, "One Last Time" | N/A | Timeline and spectrogram | 35 |
| 3.1 | Nicki Minaj, "The Night is Still Young" | 0:45-1:00 | Notation | 43 |
| 3.2 | Rihanna, "We Found Love" | 0:53-1:08 | Spectrogram | 46 |
| 3.3 | Patterns that contribute to anticipation | N/A | Notation | 49 |
| 3.4 | David Guetta, "Hey Mama" | a) 0:45-1:06 b) 1:40-2:05 c) 2:48-3:06 | Spectrogram | 52 |
| 3.5 | George Michael, "Freedom 90" | a) 1:23-1:34 b) 1:55-2:16 c) 2:16-2:38 | Notation | 58 |

| | | | | |
|------|---|---|--------------------------|-----|
| 3.6 | Guns N' Roses, "Sweet Child O' Mine" | N/A | Timeline | 60 |
| 3.7 | Guns N' Roses, "Sweet Child O' Mine" | 1:01-1:17 | Notation | 61 |
| 3.8 | Nicki Minaj, "The Night is Still Young" | 2:30-3:05 | Spectrogram | 67 |
| 3.9 | Ariana Grande, "Problem" | 0:11-0:30 | Notation | 69 |
| 3.10 | Ariana Grande, "Problem" | N/A | Timeline | 71 |
| 4.1 | Britney Spears, "Till the World Ends" | 3:07-3:23 | Spectrogram | 78 |
| 4.2 | Alesso, "Heroes" | 0:53-1:08 | Spectrogram | 79 |
| 4.3 | Tove Lo, "Talking Body" | 0:33-0:42 | Spectrogram | 80 |
| 4.4 | Tove Lo, "Talking Body" | 0:49-0:59 | Spectrogram | 81 |
| 4.5 | Tove Lo, "Talking Body" | 2:41-2:52 | Spectrogram | 82 |
| 4.6 | David Guetta, "Hey Mama" | 1:52-2:04 | Spectrogram | 84 |
| 4.7 | Daya, "Hide Away" | 0:10-0:52 | Spectrogram | 87 |
| 4.8 | Tove Lo, "Timebomb" | 0:00-1:30 | Timeline and spectrogram | 90 |
| 4.9 | Britney Spears, "I Wanna Go" | 0:35-1:01 | Spectrogram | 95 |
| 4.10 | Becky G, "Break a Sweat" | 2:25-2:52 | Spectrogram | 96 |
| 4.11 | Alesso, "Heroes" | 0:53-1:08 | Spectrogram | 97 |
| 4.12 | Jack Ü, "Where are Ü Now" | 2:34-3:07 a) 2:30-2:50 b) 2:45-3:05 | Spectrogram | 99 |
| 4.13 | Demi Lovato, "Cool for the Summer" | 0:00-1:20 | Timeline and spectrogram | 105 |
| 4.14 | Taylor Swift, "Style" | 3:45-3:57 | Spectrogram | 107 |
| 4.15 | Justin Bieber, "Beauty and a Beat" | 3:39-3:48 | Spectrogram | 108 |
| 4.16 | Becky G, "Lovin' So Hard" | 3:15-3:42 | Spectrogram | 109 |
| 4.17 | Nicki Minaj, "The Night is Still Young" | 3:25-3:47 | Spectrogram | 110 |
| 4.18 | Queen, "Crazy Little Thing Called Love" | 2:25-2:43 | Spectrogram | 111 |
| 4.19 | Ellie Goulding, "Burn" | N/A | Timeline | 112 |

| | | | | |
|------|--|------------------------------|----------------------------|-----|
| 4.20 | Ellie Goulding, “Burn” | 0:00-1:10 | Timeline and spectrogram | 113 |
| 4.21 | Ellie Goulding, “Burn” | 0:11-1:06 | Diagram | 116 |
| 4.22 | Ellie Goulding, “Burn” | 0:11-2:12 | Diagram | 117 |
| 4.23 | Ellie Goulding, “Burn” | 2:09-3:15 | Timeline and spectrogram | 119 |
| 4.24 | Ellie Goulding, “Burn” | N/A | Timeline and spectrogram | 122 |
| 4.25 | Sonic functions on two structural levels | N/A | Diagram | 124 |
| 5.1 | Full spectrograms and form timelines of 20 th -century pop and rock songs | N/A | Timelines and spectrograms | 127 |
| 5.2 | Full spectrograms and form timelines of recent pop songs | N/A | Timelines and spectrograms | 129 |
| 5.3 | Verse-Chorus form as AAB | N/A | Diagram | 133 |
| 5.4 | Ariana Grande, “One Last Time” | N/A | Timeline and spectrogram | 136 |
| 5.5 | Ariana Grande, “One Last Time,” sonic reduction. | N/A | Diagram | 138 |
| 5.6 | Taylor Swift, “I Knew You Were Trouble” | N/A | Timeline and spectrogram | 141 |
| 5.7 | Taylor Swift, “I Knew You Were Trouble” | a) 0:03-1:16 b) 0:03-0:40 | Diagram | 142 |
| 5.8 | Taylor Swift, “I Knew You Were Trouble” | 1:16-2:29 | Diagram | 143 |
| 5.9 | Katy Perry, “Unconditionally” | 0:00-1:07 | Timeline and spectrogram | 146 |
| 5.10 | Demi Lovato, “Cool for the Summer” | 2:10-2:55 | Timeline and spectrogram | 149 |
| 5.11 | Katy Perry, “Roar” | 2:34-2:58 | Timeline and spectrogram | 150 |
| 5.12 | Tove Lo, “Talking Body” | 2:10-2:55 | Timeline and spectrogram | 151 |

| | | | | |
|------|--|-----------|--|-----|
| 5.13 | Taylor Swift, “Out of the Woods” | N/A | Timeline and spectrogram | 154 |
| 5.14 | Britney Spears, “Till the World Ends” | N/A | Timeline and spectrogram | 159 |
| 5.15 | Lady Gaga, “Bad Romance” (Mark Spicer analysis) | N/A | Spectrogram and formal analysis by Mark Spicer | 162 |
| 5.16 | Lady Gaga, “Bad Romance” (Asaf Peres analysis) | N/A | Timeline and spectrogram | 164 |
| 5.17 | The Chainsmokers, “Roses” | N/A | Timeline and spectrogram | 167 |
| 5.18 | The Chainsmokers, “Roses” | N/A | Timeline | 169 |
| 5.19 | The Chainsmokers, “Roses” | N/A | Timeline | 171 |
| 5.20 | Rihanna, “We Found Love” | N/A | Timeline and spectrogram | 174 |
| 5.21 | Rihanna, “We Found Love” | 0:00-1:23 | Timeline and spectrogram | 177 |
| 5.22 | Rihanna, “We Found Love” | 2:08-3:00 | Timeline and spectrogram | 179 |
| 5.23 | Rihanna, “We Found Love” | N/A | Timeline and spectrogram | 181 |

List of Abbreviations

| | |
|------------|---------------------------|
| CHR | Contemporary Hit Radio |
| DAW | Digital Audio Workstation |
| dB | Decibel |
| EDM | Electronic Dance Music |
| EQ | Equalizer |
| Hz | Hertz |
| kHz | Kilohertz |

Abstract

This dissertation proposes a methodology for analysis of 21st-century pop music, with a particular focus on songs released between 2011 and early 2016. One of the primary factors that distinguish this music from most other popular music genres is the dominance of a sonic syntax, originating in electronic dance music (EDM), which is governed by traditionally subordinate musical features such as timbre, texture, and spatialization.

The sonic syntax that I am proposing includes three major functions: *setup*, *buildup*, and *peak*. The setup and peak functions represent the low and high points of energy in a sonic cycle. The buildup function generates an increased sense of anticipation toward a sonic peak. This is done by using techniques of varying intensity, among which are: incrementally adding sonic layers; abruptly dropping the sonic energy below the song's baseline level; intensifying rhythmic activity; and producing ascending linear gestures, such as filter sweeps and pitch bends (*glissandi*).

Methodologies designed specifically to address this genre and its technological aspects are laid out in Chapter 2, which explains central functions of the digital audio workstation (DAW) and discusses the technological skills a theorist must acquire in order to study this music. In addition, it proposes listening strategies that focus on subtle sonic details, and the use of spectrograms as visual representations of sonic textures and processes. Chapter 3 examines the

diminishing syntactical dominance of tonality in this music, in favor of the sonic syntax. In Chapter 4, the various techniques by which the buildup function is fulfilled are detailed, and Chapter 5 examines the intertwining of the sonic dimension in this music with traditional songwriting features, including lyrics and melody, in defining the sections in verse-chorus form.

Pop songs are often scrutinized according to traditional notions of musical value. This type of evaluation overlooks the way in which modern technology has revolutionized this genre. Understanding the sonic syntax illuminates a relatively new technologically based form of artistic expression that is at the core of the creation of this music, and represents a significant musical innovation.

CHAPTER 1

INTRODUCTION: SONIC SYNTAX AND ANALYSIS OF CONTEMPORARY POP MUSIC

Prior to embarking on this dissertation project, I became familiar with contemporary pop music primarily by listening to Contemporary Hit Radio (CHR) stations while driving, or by putting on my pop playlist as background to everyday activities. In short, I was a fan who listened to pop music in a very casual way, albeit more frequently than most of my classically-trained-musician friends and colleagues. Occasionally, someone would ask me about my affection for pop music, and like any respectable classically-trained composer/theorist, I would justify it by pointing to a cool chromatic alteration in Britney Spears's "Toxic" (2003) or in Beyoncé's "Single Ladies" (2008). But at some point, I began to realize that it was not the chromatic alterations (which were becoming increasingly rare) or funky chord progressions that attracted me to this music. My justification was the sound – not only the general timbre of a song, but how that sound was being used. It was a new musical language, one that had existed for a while but that I had overlooked because sound was always subordinate, in my mind, to pitch and rhythm as a creator of musical meaning. Once I let go of this prejudice, I discovered a new musical world that was nearly uncharted territory in terms of music-theoretical

understanding. I noticed that tonality generated less syntax in this music, with sound production generating more of it instead. Moreover, I could not help noticing a striking similarity between this new sonic syntax and traditional tonal syntax. Though each has its own musical objects – tonality relies on pitch relationships and the sonic syntax on timbral, textural, and spatial manipulations – they both operate on similar principles of tension and release.¹ This similarity allowed me to begin to form a new understanding of the sonic syntax in today's pop music, while relying partially on established music theoretical ideas.

What is Sonic Syntax?

I define sonic syntax as a musical grammar that relies on manipulation of timbre, sonic density (the presence and amplitude of frequencies across the sonic spectrum at any given moment), and rhythmic intensity. I refer to the sum of these elements as *sonic energy*. In most recent pop songs, a section or subsection functions as a sonic *setup*, *buildup*, or *peak*. The setup and peak respectively represent the relative low and high instants of sonic energy in a sonic progression, while the buildup represents the gradually increasing tension between these points, caused by either increasing sonic energy or by radically and abruptly decreasing it.

Sonic syntax is present, to a certain extent, in all music. It is expressed via dynamics and orchestration techniques, and in the case of studio-produced music, also through audio effects and recording techniques. However, these have traditionally been used to enhance established tonal syntax. For instance, a crescendo will often support a highly anticipatory passage leading up to a cadence, and the timbre of an electric guitar might be altered by a distortion effect upon arrival at a chorus, which was preceded by a dominant-functioning pre-chorus. Electronic dance

¹ Timbre, gesture, and spatialization were suggested by Aaron Liu-Rosenbaum as core aspects to be explored in a sonic narrative (Liu-Rosenbaum 2012).

music (EDM) was the first popular music genre to introduce sonic syntax as governing the dramatic progression of a piece of music, largely operating on buildup-drop progressions (Butler 2006). Pop music in the 21st century has gradually adopted this syntax while gradually relegating tonal syntax to a secondary role.

Background

The 1960s marked the beginning of an era in which the goals of record production were transformed from attempting to capture the experience of a live performance to creating a new sonic experience that could only be produced in the studio (Moorefield 2005, 9-34). Producers such as Phil Spector and George Martin revolutionized recorded music by treating the studio as an instrument with which they could create sounds that were impossible to achieve in a live setting.

The transition from analogue to digital recording was another big step, starting with the Soundstream platform in the 1970s (Barber 2012). The ability to digitally record and synthesize sounds has completely altered the way many producers approach the construction of a song. Digital recordings are often put together like a Lego structure. They are comprised of separate samples and computer-generated sounds that exist exclusively in digital form, are heavily processed, and are assembled on the grid of a digital audio workstation (DAW) to produce a single song. This technology laid the foundation for the sound-production-based musical language that has emerged in recent decades, first in EDM, and more recently in pop music.

The adoption of the sonic syntax described above by pop music did not occur overnight. Influences of EDM on pop can be traced back to songs such as Donna Summer's "I Feel Love" (1977), but only in the early 21st century did sonic syntax become consistently dominant in the

most popular hit songs. My own awareness of this process was triggered in 2007, when a Britney Spears album entitled *Blackout* was released. Many songs on this album, such as “Piece of Me,” “Radar,” and “Gimme More,” were produced by using samples and synthesized sounds almost exclusively, with heavy processing of the vocals to make them sound intentionally unrealistic or unnatural. Synthesizers and samplers were used by many earlier mainstream pop albums, but those still featured acoustic instruments played by musicians, and they attempted to keep the vocals sounding “natural.” *Blackout* did away with all that, and introduced an unapologetic machine-like sound, that would later influence many top-40 hit songs and albums.

In subsequently-released pop albums, I began to notice a heavier use of the EDM riser as an anticipatory process, as well sharper gaps in sonic density between sections. These developments are extremely significant - at the beginning of Chapter 5, I compare the spectrograms of three 20th-century rock and pop songs with three recent pop songs. The difference is quite striking. While there is minimal change in sonic density between sections in the 20th-century songs, the sonic alterations in the contemporary songs are so substantial that one can read the spectrograms as sonic scores that delineate the songs’ formal structures.

Literature

These recent developments in contemporary pop music have yet to receive much attention in the music theory literature. Although the body of popular music scholarship in music theory has grown substantially in the last ten to fifteen years, 21st-century commercial pop music is still little represented as a research subject in this field. Articles on popular music published in music theory journals overwhelmingly focus on 20th-century rock music, with their attention mostly directed toward form, harmony, voice leading, and rhythm. Of all the articles published between the years 2010 and 2015 in two central music theory journals, *Music Theory Spectrum*

and *Music Theory Online*, very few mention mainstream 21st-century pop music. Two articles address it in the context of mashups (Boone 2013; Adams 2015). Several others include recent pop songs as examples, but they do so in the course of discussing phenomena that apply to a wider range of popular music genres, not specifically to pop music of the recent decade. This is a significant gap in the literature, considering, not only the cultural impact of this music, but also the technological and artistic revolution that it represents.

One possible reason for this gap is the centrality of timbre as a component of sonic syntax. The scarcity of timbral analysis in popular music literature has been noted by a number of scholars, including Mark Spicer (2005) and David Blake (2012). Both cite difficulties in visual representation as a factor in the reluctance of music theorists to attempt to analyze timbre. Blake also asserts that “...timbre is especially frustrating for analytic description, at once the most apparent and least systematizable musical parameter” (par. 2.1) Simon Zagorski-Thomas criticizes the discipline more directly in a blog post entitled “How Does Music Work?” (March 31, 2011). He likens the nearly-exclusive focus on issues of “structure and harmonic progression” to his attempt to use a Beethoven piano sonata as a template for composing a piece of electronic music:

Of course, it sounds nothing like Beethoven: it uses just one feature of the Beethoven sonata and ignores all the others. In some ways, this is what I think modern musicology is doing. By not developing the language and theory to describe the gestural complexity of musical sound, by sticking to structure and harmonic progression and, more importantly, by studying the score rather than the sound, musicology is missing the point.

The primary aim of this dissertation is to develop the language and theory alluded to by Zagorski-Thomas. In its core chapters, I discuss the context in which the sonic syntax emerged; detail the various techniques by which it is expressed; and present song analyses that explore its

functional properties. Most of the timbral and textural processes I describe are clearly visualized with the help of spectrograms, particularly since the sonic traits of this music lend themselves to this type of visualization.

Why I Chose Pop and not Electronic Dance Music

As I mentioned above, the sonic syntax in pop music has largely originated in EDM. Why, then, did I choose pop as my primary object of study and not EDM? The first answer is that this is the genre that drew me to this syntax and with which I am more familiar. However, I did experience an internal dilemma as I was developing this topic regarding whether or not I should shift my focus to EDM. I decided against it, for three main reasons. First, EDM encompasses an extremely large number of subgenres. Navigating these many subgenres would have been impossible due to limitations of time and scope, and I was not interested in restricting my inquiry to a single EDM subgenre. Pop music borrows from many of these subgenres, and therefore contains a more generally applicable version of the EDM sonic language. Studying this language can become a point of departure for future inquiry of specialized subgenres. Second, I was intrigued by the interaction between the sonic syntax and the more traditional aspects of pop songs. In EDM, the sonic syntax is dominant, but in pop, it is intertwined with more traditional aspects of songwriting, particularly with the nearly-always-foregrounded vocals. This requires a different approach, which I find interesting, to both listening and analysis, as I describe in Chapters 2 and 5. Third, I simply thought that it was time to address contemporary pop music. This music is a monumental part of our culture, and influences the lives of millions, if not billions of people. It is also a popular subject of inquiry in many other disciplines, such as communications, sociology, and American studies. Providing insight into this music makes a valuable contribution to the field.

CHAPTER OVERVIEW

The progression of the three core chapters (Chapters 3-5) is characterized by a thread that starts with a discussion of the changes in musical behavior that define contemporary pop music, continues with a presentation of the various sound production techniques that are largely responsible for this change, and ends by discussing the role that these techniques play in formal structures. Below is a brief overview of each of these chapters.

Chapter 3

Chapter 3 discusses two developments that are central to the character of today's pop music – the emergence of sound production techniques as determinants of syntax and the evolution of tonality, which significantly diminishes and sometimes completely excludes the use of anticipatory mechanisms. These two developments occurred simultaneously – as sound production became more dominant in delineating form and determining the character of each section in recent pop songs, the use of cadential progressions, dissonant sonorities, and chromaticism has steadily declined. In essence, one set of syntactical tools replaced another, while retaining the latter's fundamental objects (e.g., diatonic triads, tonal centers).

I show this shift by drawing comparisons between contemporary pop, 20th-century popular music, and common-practice tonality. My primary basis for comparison is the underlying idea that drives each of these musical syntaxes – tension and release. I use David Huron's ideas put forth in his book, *Sweet Anticipation* (2006), as a point of departure for examining how each genre builds up and releases tension, which helps to illuminate the “changing of the guard” between tonality and sound production in today's pop music.

Chapter 4

The sonic syntax in contemporary pop music largely operates on three functions – setup, buildup, and peak. While the setup and the peak are defined by relative sonic density/energy, the buildup function is the driving force in this music and is represented by a wide range of techniques. In Chapter 4, I detail the most common techniques for fulfilling the buildup function, and discuss their relative level of intensity and structural significance. Identifying these techniques is particularly important in understanding complex sonic structures, as the beginning and end of the buildup define the borders of the other two functions, on multiple structural levels.

Chapter 5

While the functions described above play a significant role in creating the formal structure of pop songs, the core song sections (i.e., verse and chorus) are still largely characterized by their traditional defining aspects, lyrics and melody. Chapter 5 explores the intertwining of the traditional and sonic dimensions in today's pop. I begin this chapter with examples of songs in which the traditional and sonic dimensions correlate (i.e., the verse acts as a sonic setup, the pre-chorus as a buildup, and the chorus as a peak), and proceed to discuss songs in which they diverge and forge complex relationships. Additionally, since a large portion of recent pop songs feature a single unchanging chord loop, I discuss the typical sonic traits of sections that were previously defined by harmonic function, such as the post-chorus, bridge, and transitions. Finally, I present analyses of two songs that pose interpretive challenges – “Roses” by The Chainsmokers (2015) and Rihanna's “We Found Love” (2011) – and provide a more in-depth discussion of my decision-making process when approaching such songs.

Technological Tools

Although I use some traditional notation in Chapter 3, the vast majority of my analytical work required the use of additional technological tools in order to address the sonic properties of this music. The primary tools I used were *Ableton Live*, *Sonic Visualizer*, and *Variations Audio Timeliner*.

Ableton Live is a digital audio workstation. It was developed in Germany by Gerhard Behles, Robert Henke, and Bernd Roggendorf. I used it to reverse engineer (attempt to recreate) sounds from pop songs, as well as to explore various techniques for achieving certain sounds and effects. In chapter 2, I use screenshots from its interface to show some of its capabilities, which are central to the work of today's pop producers. *Sonic Visualizer* was developed at the Centre for Digital Music at Queen Mary University of London. I use it to generate spectrograms, which visualize sonic activity over the course of a song. Finally, *Variations Audio Timeliner* was developed by the Indiana University Digital Library Program (specifically by Brent Yorgason, Jim Halliday, and Chris Colvard, with conceptual guidance from Eric Isaacson, according to the official website). This tool provides a platform for generating timeline analyses of the formal structure of a song (or any musical piece). I frequently use these timelines in conjunction with spectrograms, to show the correspondence of specific sonic activity to song sections.

A note to the readers: The analyses presented in this document all refer to the official audio versions of the songs, as detailed in the discography. If you choose to listen to the recordings via music videos on YouTube or another video platform, please be aware that some music videos add extra footage before, after, or in the middle of a song. Therefore, the timings of

events to which I point in the analyses may differ from the official versions, and your reading of the timeline should adjust accordingly.

CHAPTER 2:

CHALLENGES AND METHODOLOGY

Analyzing a currently-evolving genre such as contemporary pop music presented several challenges. First, the existing literature addressing this music and its unique characteristics is very limited, and therefore there were no established analytical methods specific to this genre upon which I could rely. Second, while I could transcribe the melodies and chord progressions of these songs, a traditional score could not represent the aspects in which I was most interested, such as gradual timbral alterations and the amplitude (loudness) of certain frequency bands within the sonic spectrum at given moments. This meant that I needed to acquire certain technical skills, including mastering a digital audio workstation (DAW) and reading spectrograms, for example. Third, pop music trends continuously changed over the course of my dissertation work, which made it more difficult to arrive at conclusions regarding the behavior of this music, but also helped me to better understand the patterns that I was able to identify.

This chapter details how I have addressed those challenges. I will discuss my song selection process, the technological aspects of this dissertation, the listening strategies that I

employed as part of my analytical approach, and the ways in which I represent musical examples visually.

Accessibility to Music Theorists and Pop Practitioners

Although this is a music theory dissertation, one of its aims is to explore a common ground between ideas generated by the scholarly music theory discipline and those generated and put into practice by songwriters, producers, and sound engineers who work in the popular music industry. To make this document accessible to both worlds, I explain concepts that may be familiar to some readers but not to others, depending on the background of the individual reader. For instance, trained music theorists are quite familiar with the ideas of stability and instability, as traditionally framed in theories of tonal music. Likewise, every audio engineer has intimate knowledge of the equalizer (EQ) and what it does. However, elaborating both the notion of stability/instability and the function of an EQ in the body of this dissertation (rather than in footnotes or references) is essential to the construction of ideas and methods that are central to it.

GENERAL CHALLENGES RELATED TO ANALYZING CONTEMPORARY POP

Developing the topics of the three core chapters presented different challenges and each required its own methodology, but before I could solidify these narrower topics, I had to address more general challenges related to the nature of my objects of study. First, I needed to address the question of what constitutes contemporary pop music. Since these songs are extremely recent, I did not have a historical perspective from which to position songs within a genre. Thus, I had to listen to a large number of songs, identify patterns that characterize this genre, and narrow the sample down to a group of songs that exhibit these patterns. In addition, as mentioned

above, this research also required technological skills beyond those normally expected from a music theorist, as well as listening strategies that would help me to address both the traditional and sound production aspects of this music. In the following sections, I will detail the ways in which I dealt with each of these challenges.

Selecting Songs for Analysis

One aspect of writing a dissertation about a currently-evolving musical genre is that songs for analysis cannot be chosen at the beginning of the process, but must be added to the sample as they are released throughout the entire process. This also means that the examined features of this music change while the dissertation is being written. For instance, while in 2011 many of the top hits focused exclusively on synthesized sounds, more recent hits tend to integrate instruments such as guitar and saxophone in the electronically produced sonic environment. This genre has also evolved in other ways, including heavier use of manipulated audio samples rather than purely synthesized sounds, and more experimentation with form as a result of the continually-increasing influence of EDM, rap, and indie music.

I also had to determine which songs qualify for my study. Most genres are defined by specific sonic traits. We can distinguish a Romantic symphony from a Classical one by the size and instrumentation of the orchestra and the harmonic language. Speed metal is characterized by heavily distorted guitars playing extremely fast solos over ferocious drum and bass parts. In jazz, the rhythm section and solo instruments (e.g., saxophones and trumpets) are ubiquitous, as are the numerous melodic and harmonic idioms that form its improvisational language. Even rock music, which encompasses several sub-genres, can be identified by a general sound that relies on guitars, drums, bass, and sometimes keyboards, the timbre of the vocalist's voice being a key variable in creating a band's specific sound. Pop music is much more flexible from a sonic

standpoint, and often borrows from several genres with more defined sounds, treating them as musical ingredients in a new sonic mix. For example, Meghan Trainor's "All About That Bass" (2014) and Ellie Goulding's "Burn" (2013) are both pop songs, though they differ greatly in their general sound. In fact, since today's pop stars are typically branded as individual singers rather than integral parts of bands, they are not bound to specific players and instruments. Thus their sound, which is normally electronically produced (aside from the vocals) and occasionally supplemented by hired instrumentalists, can significantly vary from song to song. Some might define pop music as "whatever happens to be popular at the moment," but this would make explaining the various periods in the last half-century, when rock bands such as Led Zeppelin, Kiss, Nirvana, and Guns N' Roses ruled the charts, extremely difficult.

Some features can certainly be found in most pop songs. The length of a song is usually somewhere between three-and-a-half and four minutes. Most pop songs use some variation of verse-chorus form (verse-chorus; verse-chorus; bridge-chorus). A four-chord cycle, repeated throughout the song, is also typical of this genre. The vocal melodies and rhyme schemes contrast between sections. But each one of these traits can be associated with other genres of popular music that are not necessarily pop. On their own, they cannot act as identifiers of the genre. For my sample to represent the truly unique musical contributions made by recent pop music, I decided to include only songs whose sounds were constructed using sound design and other sonic manipulations in ways that were not ubiquitous in past music. So, although "All About That Bass" was a major pop hit in 2014, it did not make it into my sample, because it does not contain the specific emerging traits on which my research is focused.

The first few songs in my sample came from my personal playlist. I found the use of sonic sweeps in Britney Spears's "Till the World Ends" (2011), the sound design in Kendrick

Lamar's "Swimming Pools" (2012), and the trajectory of sonic energy in Rihanna's "We Found Love" (2011) very interesting and worth exploring through an analytical lens. After focusing on these songs and a few others for a period of time, I expanded my sample to include new songs that were featured on Contemporary Hit Radio stations, and whose sound grabbed my ears. I also periodically scanned the main pop charts in order to listen to newly-released songs, and added those in which the sonic syntax was dominant to my sample. I did not initially set out to analyze a specific number of songs, but as my sample was nearing a hundred songs, I felt that I was reaching saturation in terms of exploration of unique themes, and decided to cap my sample at that number, occasionally replacing some songs with newer ones to better represent evolving trends. The vast majority of the songs that I have analyzed for this dissertation were released between the years 2011 and 2016, with only a handful of songs having been released in earlier years.

Establishing Analytical Methods

Music theory literature has thus far treated contemporary pop music as an extension of 20th-century popular music. Mentions of recent pop songs have been made in the context of larger topics related to popular music as a whole, rather than as an exploration of the unique traits of this music. Nevertheless, the existing literature contributed to my understanding of some aspects of my topic. I was able to rely on it for established ideas about harmony and form in 20th-century popular music, put forth by Mark Spicer (2011), Drew Nobile (2015), David Temperley (2011), Walter Everett (1999), and Jay Summach (2011), among others. Moreover, some relevant aspects of EDM have been examined by scholars such as Mark Butler (2005) and Ragnhild Torvanger Solberg (2014), particularly the buildup-drop sequence, which I will detail below. Otherwise, to the best of my knowledge, music theory literature specifically addressing

contemporary pop music, in which the sonic syntax originating in EDM and traditional songwriting syntax have become completely intertwined, is extremely limited.

Thus, I had to analyze a large number of songs from every perspective I could think of. I examined the formal structure of every song in my sample. I made note of as many subtle gestures as I could find and explored ways in which they participate in sonic themes. I attempted to reverse-engineer many songs in order to better understand the makeup of their sounds and how they relate to one another. In short, I tried out as many methods as possible to see which ones would yield meaningful insights about this underexplored musical language.

Technology

The sonic dimension central in today's pop music requires a different type of musical training than the theorist traditionally receives – technological musicianship.² In order to understand this music, one must become familiar with the tools used to create it. The primary technological environment in which this music is created is the digital audio workstation (DAW). Modern DAWs are extremely complex apparatuses, and I cannot say that I have mastered my DAW, *Ableton Live*, to the level of a professional audio engineer. However, there are some DAW functions that are used to shape the sound and syntax of nearly every modern pop song, and these were crucial for me (as well as for readers of this dissertation) to understand.

² A recent debate between Ian Pace and Simon Zagorski-Thomas, posted by Pace on his blog (April 27, 2016), reveals a misconception held by many regarding popular music. Pace refers to the classical music tradition as “literate,” implying that the practice of popular music does not require a high level of literacy. He does not seem to value the technological literacy required to produce professional-quality pop songs, or perhaps he is not aware of the complexities of this type of literacy. For the full exchange, see <https://ianpace.wordpress.com/2016/04/27/responses-to-simon-zagorski-thomass-talk-on-dead-white-composers/>

Equalization (EQ)

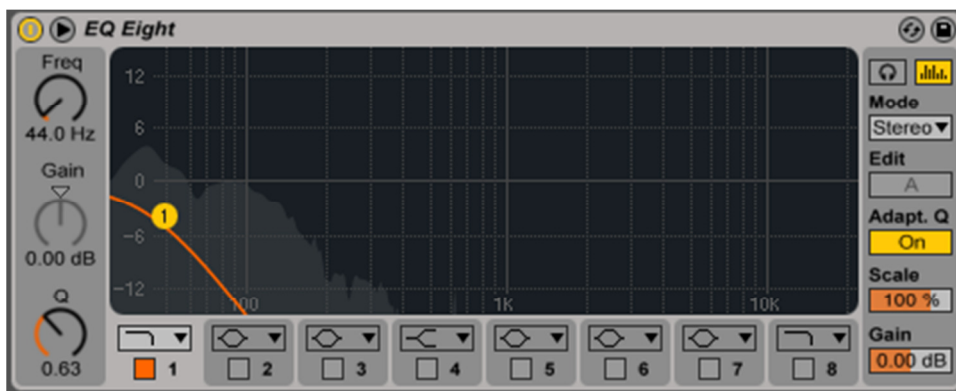
The equalizer is the effect that is most widely used by record producers today. It alters the amplitude (loudness) of frequency bands defined by the user. Different equalizers serve different purposes. Some manipulate broad ranges of frequency bands and some target more specific ones. Manipulating the sonic spectrum in this way can, for example, create the illusion that a voice is sounded through a phone or that an orchestral piece is being played by an old AM radio. It can also be used to sonically shape a sound in order to make it more focused and to carve out sonic space for another sound so that the two do not clash.

One way in which EQs carve out sonic space is by acting as sonic filters. A high-pass (or low cut) filter cuts or mutes all frequencies below a user-determined threshold. Inversely, a low-pass (or high-cut) filter, muffles or mutes frequencies above the determined threshold. Other filters, such as shelving or peaking filters, allow their user to boost or muffle selected frequency bands. The possibilities are nearly endless.

EQs have been an indispensable part of popular music record production since its earliest days. However, one way in which they have become a defining resource of modern-day pop is the use of the *filter sweep*. Commonly, pop producers use a low-pass filter, initially cutting out all but the very low frequencies in the spectrum, and then gradually increasing the cutoff frequency. Example 2.1 shows three points along a filter sweep: the beginning, middle, and end. The orange line represents the EQ's function at every moment (the number of decibels per octave that are reduced or gained), and the lightened gray area represents the sonic activity (the amplitude of each frequency in the spectrum at a given moment). In Example 2.1.a, the low frequencies are progressively cut until they are almost completely muted at around 500 Hz. The

controller is then shifted so that the frequencies that were previously muffled are gradually re-amplified and the frequencies that were completely muted are now more and more present, until the entire spectrum is “released” and the highest frequencies are boosted, causing the mix to sound much brighter than it would without the EQ treatment. This produces a sweeping timbral transformation, and can also function as a crescendo when applied to the entire mix or a significant part of it.

a) **Beginning.** The low-pass filter significantly reduces the amplitude of frequencies up to ~400 Hz, and completely mutes frequencies beyond that range. The music sounds soft and muffled, with low-range instruments, such as the bass and kick drum, having the most presence in the mix.



Example 2.1. *Stages of a typical filter sweep* (screenshots from Ableton Live’s Operator synthesizer). The horizontal axis represents frequency (Hertz). The vertical axis represents amplitude (decibels). The orange line indicates the number of decibels added or cut to frequencies in the sonic spectrum. The light gray area indicates the relative amplitude of frequencies. The filter is gradually moved in a sweeping motion, making the overall sound progressively brighter (and in this case, louder).

b) **Middle.** After being gradually released, frequencies up to 150 Hz regain their original amplitude; frequencies between ~150 and ~900 Hz are slightly amplified; frequencies between ~900 and ~6,500 Hz are now audible but their amplitude is still reduced; and frequencies beyond ~6,500 are still muted. The sound becomes more balanced, with the vocals and other instruments gaining presence in the mix but still sounding somewhat muffled.



c) **End.** Most of the audible frequencies regain their original amplitude. Frequencies above ~2 kHz are amplified, which makes the overall sound particularly bright (despite the cut beyond ~16 kHz).



Example 2.1. *Cont.*

Filter sweeps are often the centerpieces of *risers*, which build up a heightened sense of anticipation in the listener, whose attention is directed toward the rising frequencies until it is time for the *drop*. The drop is the arrival at a climactic passage in which the lower frequencies,

normally represented by the kick drum and bass, regain their prominent status in the mix, after being temporarily removed or obscured. This is known as a *buildup-drop*, or *riser-drop* sequence. An example of a riser-drop sequence, to which I will refer several times, can be heard in “Till the World Ends” at 3:07.

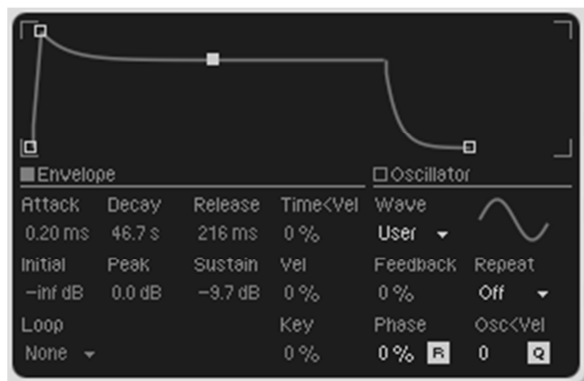
The filter sweep can also be used to transform the timbral character of a sound, such as the guitar sound in the opening ten seconds of Taylor Swift’s “Style” (2015), which is transformed from an electric guitar sound with an overdrive/distortion effect into a clean acoustic guitar sound. Other functions of the filter sweep include moving sounds between the background and the foreground of the mix, and transforming a sound’s sonic depth and range. These functions will be discussed throughout the analyses in this chapter.

In sum, the EQ plays a major role in the timbre and spatial positioning of each sound in a song. By applying EQ settings to each track in the mix, the producer and sound engineer can design a unique sonic identity for a song and ensure that all of the sounds fit together in the sonic space without clashing. Through gradual manipulations, such as the filter sweep, they are also able to use the EQ to create dynamic elements, such as anticipatory gestures (e.g. risers), spatial motion, and timbral transformation.

Timbral Shaping

In addition to the EQ, the DAW provides a large number of tools to determine and manipulate the timbral qualities of each sound. Producers can start with a computer-generated sine wave and apply various processes to transform it into almost any sound. When creating a synthesized sound, one has complete control over the two most important factors in determining the timbre of a note – the *envelope* and the *amplitude ratio of the harmonic partials*.

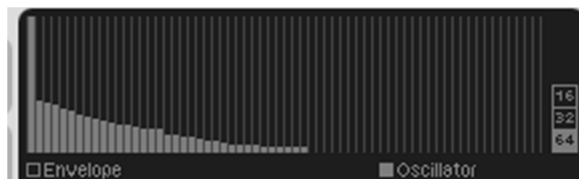
Example 2.2 shows the parameters for manipulating the envelope of a pitch in the *Ableton Live* DAW. The attack is quite fast – 0.20 ms, which means the peak volume of a note triggered under these settings is reached almost instantly. The sustain – the volume maintained after the initial decay – is 9.7 dB below the peak, but since the decay is quite slow, 46.7 seconds, we will not reach that level unless the note is held at least as long as the decay time. The release time is a moderate 216 ms. Setting the attack time at a slower rate will cause the sound to resemble a clarinet. Extending the release time will give it a characteristic similar to that of a bell, while a very short decay time combined with a sharp decline in the amplitude of the sustain parameter will transform the sound into a pitched click.



Example 2.2. *Envelope settings in Ableton Live's Operator synthesizer.*

Example 2.3 shows the platform for manipulating the amplitude ratio of the harmonic partials in the same DAW. The leftmost bar represents the fundamental tone, and each slot from left to right represents the next partial in the harmonic series. Acoustic instruments naturally produce harmonic partials, and the variation in the amplitude ratio of these partials relative to the

fundamental tone plays a significant role in the differences in timbre between one instrument and another. DAWs allow producers to begin with an artificially produced sine tone with no harmonic partials, to which they can add partials and manipulate their amplitude in order to control the tone's timbre. If a partial is given the same amplitude as the fundamental tone, it is audible as an independent pitch. If the amplitude is low enough so that it is masked by the fundamental tone, its function is to alter the tone's timbre to create a more complex sound.



Example 2.3. *Harmonic partials settings in Ableton Live's Operator synthesizer. Each vertical line, from left to right, represents the next partial in the harmonic series (the first line is the fundamental tone). The height of the lines represents relative amplitude.*

Another way to arrive at new sounds is to manipulate recorded samples. DAWs enable the producer to transpose, slow down, speed up, and otherwise deform an original sample to the point that it becomes unrecognizable and is transformed into a new sound. Skrillex and Diplo, two prominent EDM producers who have collaborated with mainstream pop artists as part of the growing trend of EDM-influenced pop, comment on this in a New York Times video feature about their 2015 hit “Where Are Ü Now” (featuring Justin Bieber).³ They discuss manipulation of audio samples as a way to achieve new sounds that cannot be achieved by pure synthesis, and

³ Pareles, Jon. “Bieber, Diplo, & Skrillex Make a Hit.” Nytimes.com Video, 7:59. Posted August 25, 2015. <http://www.nytimes.com/interactive/2015/08/25/arts/music/justin-bieber-diplo-skrillex-make-a-hit-song.html>

reveal that the primary melodic hook in the song, which was described by some as “the violin-flute sound,” is actually a heavily processed sample of Bieber’s vocals.

Of course, controlling the envelope and the amplitude ratio of the harmonic partials or manipulating audio samples are just the tip of the iceberg. Professional producers and recording engineers routinely create elaborate chains of processing that produce extremely complex sonic sequences. One can combine manipulations of the envelope and the amplitude ratio of the harmonic partials, apply them to individual notes, automate the manipulations so that many processes can occur simultaneously, and apply filters, equalizers, reverb, and other effects. It is common for a contemporary pop song to consist of 150-200 tracks in the DAW. A single kick drum sound can be the result of a combination of four or five different sounds that were all manipulated and combined to create the desired sound, as Dr. Luke demonstrates in the 2011 ASCAP Expo.⁴ The means for manipulating timbre are endless. Even timbre-oriented composers with huge symphonic orchestras at their disposal, such as Debussy, Ravel, or Berlioz, did not have the flexibility that today’s amateur producer has. As a result, timbre has become the focal point for many producers of contemporary pop music.

Panning and Spatial Positioning

Spatialization refers to positioning of sounds in the sonic space, causing the listener to perceive them as high or low, close or far away, or in various peripheral positions. It contains several dimensions that may seem unrelated: panning, volume, and pitch/frequency range.⁵

While panning causes the sound to physically come at the listener from his/her left or right,

⁴ "Dr. Luke Shares Music and Production Tips at ASCAP 'I Create Music' EXPO." YouTube Video, 3:20. Posted September 19, 2011. <https://www.youtube.com/watch?v=AWVJkokLbzQ>

⁵ Ruth Dockwray and Allan F. Moore (2010) consider time to be the fourth dimension in the “sound box,” which is a visual representation of sonic space.

increasing or decreasing the volume of a sound creates the sonic illusion of it growing physically nearer or further away, and playing higher pitches or increasing the volume of their higher harmonic partials creates the sonic illusion that they are physically higher. Nevertheless, these factors are intertwined. In an imaginary three-dimensional sonic space, a sound's volume relative to other sounds is responsible for our perception of sonic distance or depth, while pitch placement and harmonic content are responsible for our perception of sonic altitude. In context, these roles extend even further. For instance, transposing a pitch upwards (or boosting its higher harmonic partials) can move it to the foreground, even if its actual volume in decibels remains lower than other sounds, because its fundamental tone and/or harmonic content are in a less obscured position in the sonic space.

Similarly to the EQ, spatialization can contribute to the clarity of the mix by positioning the different sounds in a manner that prevents sonic clashing. It can also be used to create dynamic processes. Panning, for instance, can be used to cause the listener to perceive a sound as if it were moving around him/her by gradually shifting it from one side to the other. An example of this can be heard in Daft Punk's "Lose Yourself to Dance" (2013) at 1:36-2:52. In Ludacris and Kelly Rowland's "Representin'" (2012), the synthesized bass moves from the background to the foreground simply by gradually boosting its volume throughout the verse (0:15-0:43). Filter sweeps move a sound up or down in our spatial perception. These dynamic spatial processes are integral parts of the language of contemporary pop music.

Automation

All of the dynamic processes described above, as well as many others, are achieved via *automation*. The modern DAW allows us to draw any process we like into a track, which will then be automatically executed during playback. For instance, if I want a vocal track to "circle"

the listener, such as in the “Lose Yourself to Dance” example cited above, I can draw a line that will instruct the DAW to automatically move the panning back and forth within a timeframe that I determine. This is an invaluable tool because it enables a single sound to simultaneously undergo multiple dynamic processes, which would be impossible to do in a traditionally instrumental performance. Automation can be used for large continuous processes such as filter sweeps and fade-ins/fade-outs, or for subtle details, such as removing unwanted breathing sounds from a vocal track, or momentarily applying an EQ setting to a sound so that it doesn’t clash with another.⁶

In sum, the tools described in this section (EQ, timbral shaping, panning, and automation) are central to the work of the producer, not only from a technical standpoint, but also from the perspective of artistic approach. The methodology presented in this dissertation addresses the use of dynamic processes such as filter sweeps and automated pitch bends similarly to the way that methodologies focused on more traditional music treat harmonic functions. This type of analytical focus on sound production is important because it addresses this music on its own terms. The language used by producers to describe these tools and processes, with terms such as “buildup-drop” and “risers,” is different from how music theorists typically talk about music. This is understandable, since the music and its functional tools are different. But understanding the underlying similarities between these two musical languages can contribute to a deeper understanding of each, which is one of the goals of this dissertation.

⁶ For more information on automation, see Sam Inglis, “Creative Mix Automation in Your DAW,” <http://www.soundonsound.com/sos/aug11/articles/mix-automation.htm>

Studying Sonic Details

Despite the heavy reliance on sound production, pop music still foregrounds pitch and rhythm. Focusing on these two parameters makes it easier to hook listeners on a song with repetitive beats, melodies, and chord progressions. It is more challenging to consciously notice the sonic details that characterize a song. In order to make sure that I do not inadvertently ignore important sonic details, I used a system in which I deliberately shifted back and forth between two modes of listening — foreground listening and background/close listening. I had specific goals associated with each of these modes. In addition, becoming intimately familiar with the internal makeup of songs in this genre required engaging in reverse engineering, in which I attempted to recreate the sounds of my objects of study. I detail below the principles of each of these activities.

Foreground Listening

This mode entails repeatedly listening to a song in order to become familiar with its foreground features, such as the chord progression, melody, drum patterns, and major changes in sonic texture. Using this mode, I can map out the sections and subsections of the song, and evaluate their functions in terms of stability and sonic energy. I can take in the general sound of the song, and note specific foregrounded sonic gestures, such as risers and abrupt removal of sound layers. I can begin to examine the role of the pitch material and lyrics in the song. In short, foreground listening entails paying attention to everything that a casual listener would consciously hear. But while casual listeners often think of the melody, harmony, and rhythm as “the music,” and sound production as a secondary supporting element, this mode of listening treats sound production as an equal, and often more important, aspect of the song’s character.

Close Listening

This mode focuses on all possible middleground and background details. These may include subtle samples, textural elements, background vocals, percussive sounds, and anything else that a listener might not consciously perceive during foreground listening. The most subtle gestures can participate in sonic themes. A background sonic layer can help to shape the sonic trajectory of a song by increasing the sonic density of a third chorus compared to the second. At times, important structural processes can only be noticed if the listener diminishes his/her focus on the foreground. For instance, in Ellie Goulding's "Burn" (2012), which I analyze in detail in Chapter 4, an extremely subtle process that gradually brightens the texture over the span of several sections has significant implications for my listening experience and for my interpretation of the dynamic trajectory of the song. This important process would have likely remained in my subconscious had I not engaged in this mode of listening.

At another point, close listening enabled me to hear the details of sonic treatment in Lady Gaga's "Bad Romance" (2009). Mark Spicer identifies four iterations of a section that he interprets as the post-chorus.⁷ Focusing on the song's sonic dimension, I noticed different sonic treatment in each of these iterations, though the second and third iterations were nearly identical (these alternatives are discussed in Chapter 5). This difference is extremely subtle and unlikely to be detected by a listener focused on the foreground, but it transforms the character of this section and contributes to the sonic variety of the song.

To clarify, noticing these subtle details and taking them into account does not mean that one is privileging the discovery of insignificant details over the actual listening experience.

⁷ Mark Spicer, "(Per)Form in(g) Rock: A Response," *Music Theory Online* 17:3 (2011): par. 10.

These details affect the listening experience whether the listener is consciously aware of them or not. Making an effort to bring their presence into our conscious understanding of a song serves to enhance our understanding of the unique properties of that song, and to enhance our listening experience as well.

Reverse Engineering

A large portion of the time I spent studying this music was dedicated to reverse engineering – attempting to recreate the complex sounds and textures that I noticed during the close listening phase. This, in essence, is an extension of the close listening mode – after discovering the subtle sounds and processes in the music, reverse engineering allowed me to explore their internal makeup, as well as relationships between different sounds and sonic themes. Doing this is like walking in the footsteps of the producer and sound engineer, and witnessing their choices. Particularly when examining spatialization and sonic themes, reverse engineering helps in understanding which sounds and effects are used and why, as well as how they fit into a sonic narrative.

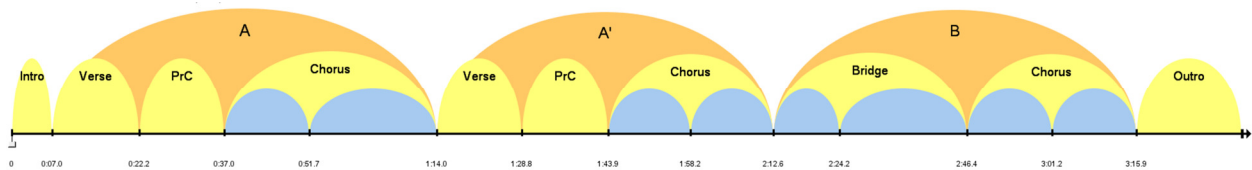
VISUAL REPRESENTATIONS

The vast majority of musical examples in this dissertation are not represented in musical notation. Rather, I use spectrograms and form timelines generated with the aid of the *Sonic Visualizer* and *Variations Audio Timeliner* software, respectively. Using these tools, I am able to visually represent the aspects of this music that are most relevant to this dissertation. The following is a brief overview of these visual tools.

Form Timelines

In addition to the traditional defining traits of sections in pop songs (e.g., verse, chorus), sections and subsections are also defined by changes in sonic texture. Thus, it is helpful for me to represent the form of a song as a timeline graph of a specific recording, allowing a relatively condensed formal overview of an entire song as a single image, and simplifying reference to sections in the song. This facilitates discussion of the relation between sections when musical notation is not possible. It also makes it easy to compare songs and to identify formal trends within the genre, as well as similarities of typical formal structures between genres.

Example 2.4 shows a formal breakdown of Ariana Grande’s “Break Free” (2014). Each section is represented by a yellow bubble, and those are grouped by orange bubbles to represent the larger-scale structure of the song. Some sections, which are longer and feature internal sonic changes, are subdivided using blue bubbles to indicate these changes.



Example 2.4. Ariana Grande, “Break Free” (2014), featuring Zedd. Form timeline.

Spectrograms

Spectrograms visualize the sonic activity in an audio file over time. In particular, they show the relative amplitude of frequencies at any given time, and can therefore illustrate changes

in sonic density, the relative brightness of certain sections, and gradual processes, such as filter sweeps, among other things. I use two types of spectrograms to visualize sonic activity – the *plain spectrogram* and the *melodic range spectrogram*. In both types, the horizontal axis represents time (in minutes:seconds), the vertical axis represents frequency in Hz, and the color brightness at each point represents the relative amplitude of each frequency at each time point. The difference between the two types of spectrograms is in the range of frequencies that they cover, and in the type of sonic activity that they show.

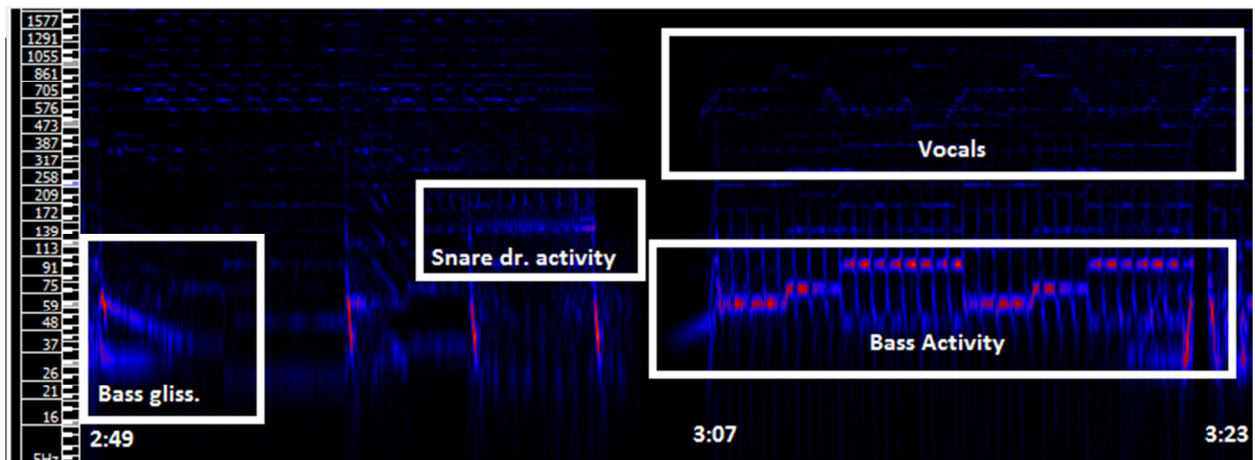
The plain spectrogram covers more or less the entire audible range of frequencies. As certain frequencies gain amplitude (i.e., become louder), they are represented by brighter colors on the spectrogram, changing from dark green to bright green, then yellow, and finally orange and red. I use the plain spectrogram to show overall changes in sonic density between and within sections, the relative presence of certain frequency bands in a given sonic texture, and processes such as filter sweeps and other notable timbral transformations.

The melodic range spectrogram represents frequencies up to just under 2 kHz (around B6), in order to focus on the fundamental tones of the sounds in the mix. It is easier to pick out the sonic activity of single sounds or tracks by using the melodic range spectrogram, because it has a much higher amplitude threshold – frequencies below a distinctly-audible amplitude level are not represented in the melodic range spectrogram. The color scheme of the melodic range spectrogram includes blue (lowest amplitude above the threshold), red, and yellow (highest amplitude), as well as a black background that encompasses all of the frequencies that did not reach the amplitude threshold. I use the melodic range spectrogram primarily as a pseudo-score, representing the activity of individual instruments/sounds along with their presence in specific regions of the sonic spectrum. For example, vertical lines in the low register that resemble flames

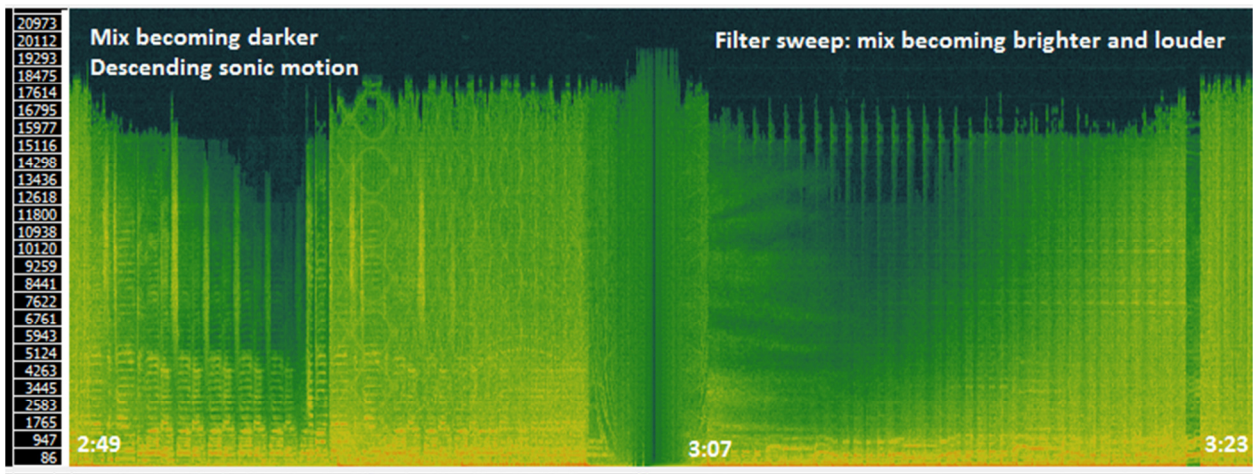
often represent loud kick drum hits, while horizontal lines in higher registers can represent melodic material.

Example 2.5 shows both a melodic range and plain spectrogram of the bridge (2:49-3:23) from the aforementioned “Till the World Ends.” In the melodic range spectrogram (Example 2.5.a) we can clearly see a downwards bass glissando at the beginning of the section, heightened snare drum activity at around 3:00, and the bass and vocals between 3:07 and 3:23 (notice the keyboard diagram to the right of the frequency scale, which shows the correspondence of notes to frequencies). By contrast, it is difficult to see these details in the plain spectrogram illustrated in Example 2.5.b, but much easier to see the general timbre of the mix become darker after 2:49, and then gradually brighter and louder over the course of a filter sweep riser, starting at 3:07. In both spectrograms, it is easy to see a pause in sonic activity just before the drop at 3:23, as all the sounds are muted, leaving only the vocals as a pickup to the final chorus.

a) **Melodic Range Spectrogram.** Activity of individual sounds/instruments is visible, due to the more limited frequency range and higher amplitude threshold for visibility.



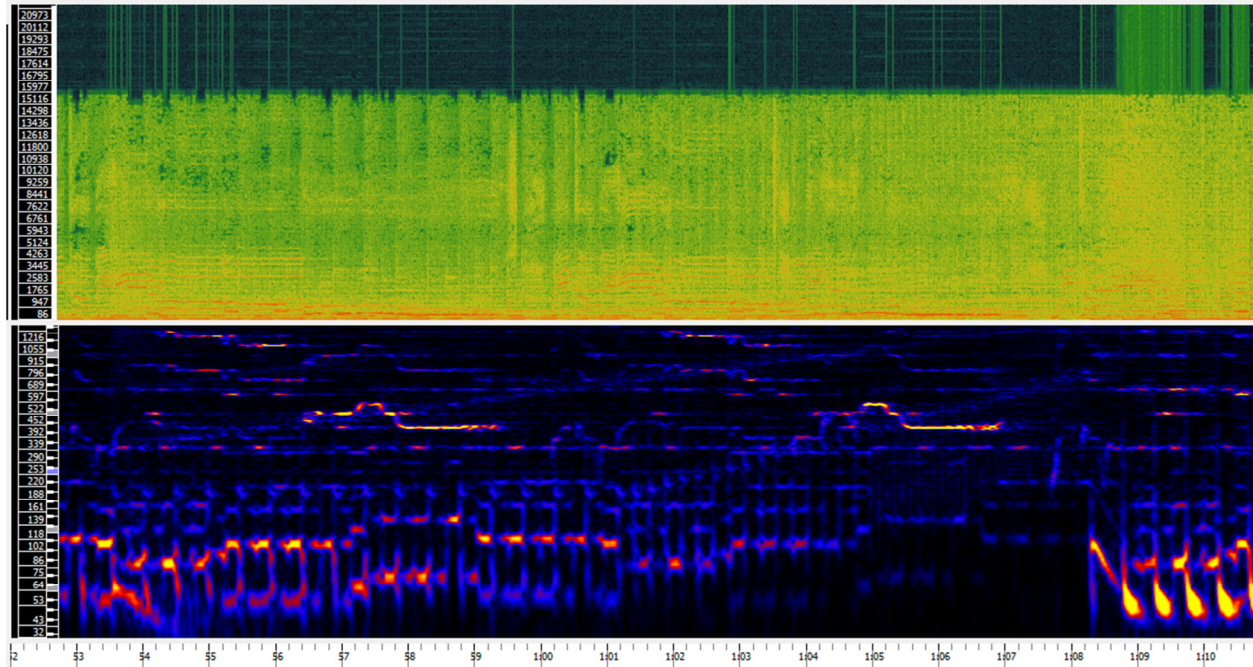
b) **Plain Spectrogram.** Large scale timbral changes and processes are visible. The entire audible frequency spectrum is included, and the lower amplitude threshold allows harmonic partials, which are not heard as independent sounds but affect the timbre, to be represented.



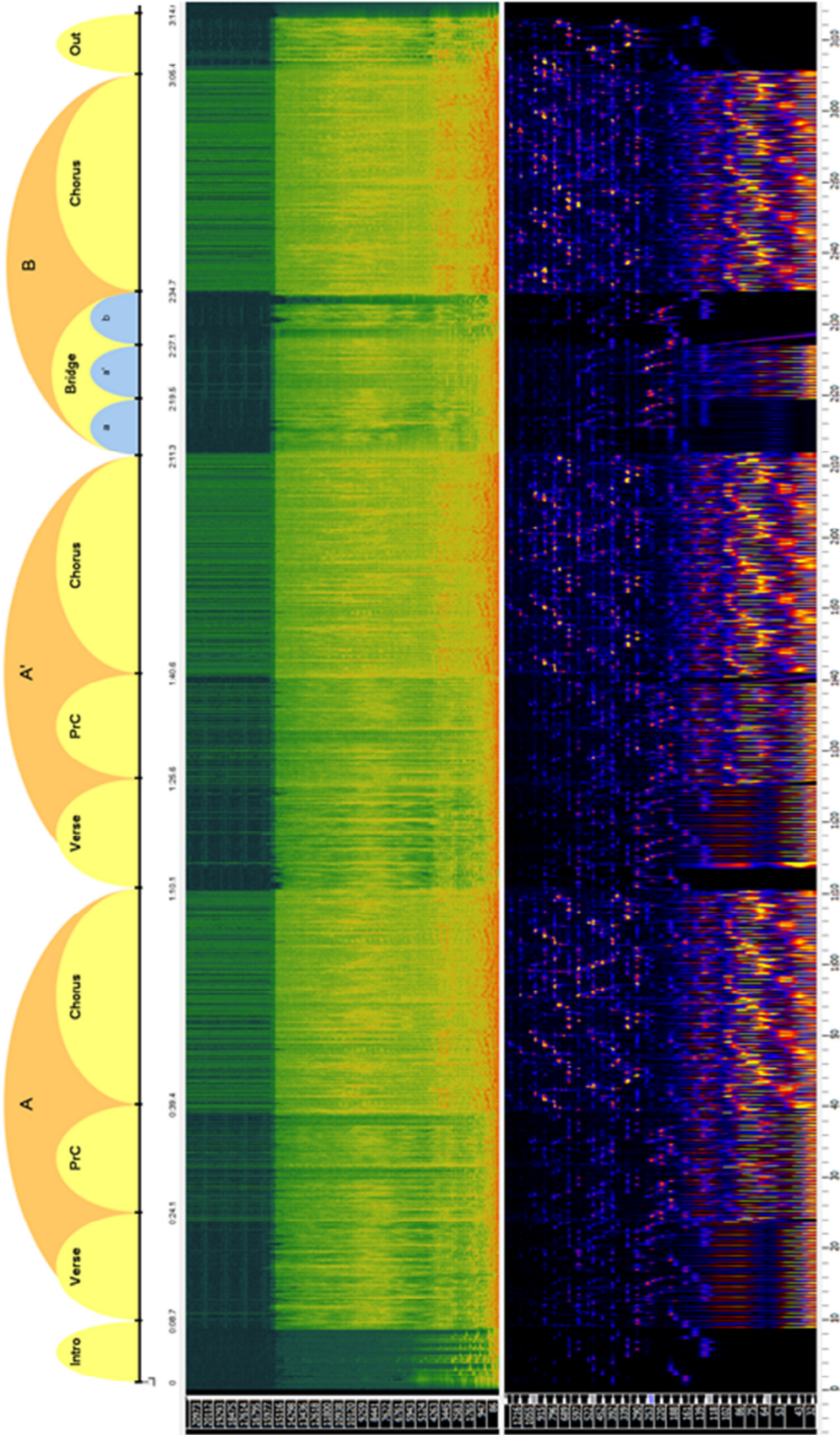
Example 2.5. Britney Spears, “Till the World Ends,” (2011), 2:49-3:23. Functions of the plain and melodic range spectrograms.

How I Use Spectrograms and Form Timelines

On occasion, I use each of the above methods of representation separately. I use the form timeline as a reference in discussions that include multiple sections of a song. I use one type of spectrogram or the other when I want to show a very specific detail, though most of my examples include both the plain and melodic range spectrograms positioned above a time axis. In many cases they are combined with a form timeline. For instance, Example 2.6 shows a riser-drop sequence from Alesso's "Heroes" (2014), which I discuss in Chapter 4. Here, I did not include a form timeline because the process only spans a single section. By contrast, when discussing processes that occur over a number of sections or an entire song, I position a form timeline above the spectrograms, showing the correlation between the section changes and sonic processes. Example 2.7 shows such a representation of Ariana Grande's "One Last Time" (2014), taken from Chapter 5.



Example 2.6. Alesso, “Heroes,” (2014), 0:53-1:10. High frequencies gain more amplitude (as indicated by the colors becoming progressively brighter in the plain spectrogram,) while the kick drum and bass are gradually obscured (as indicated by the steadily reduced presence of “flames” in the low register of the melodic range spectrogram). This forms a riser, which leads to a drop at ~1:08, represented by the return of the flames (in the MR spectrogram) and bright yellow color (in the plain spectrogram).



Example 2.7. Ariana Grande, "One Last Time" (2014). Form timeline and corresponding spectrograms.

SUMMARY

Most of the methods presented in this chapter were born of a growing awareness of the syntactical dominance of sound production in contemporary pop music, as well as from the need to incorporate technology when analyzing this music. Music theory scholars have alluded to the challenges theorists have faced in their attempts to analyze timbre (Spicer 2005, par. 14; Blake 2012, par. 2.1). However, modern technology allows record producers nearly-unlimited control over timbre, which they have used to create new forms of musical expression, and the increasing user-friendliness of DAWs and sonic visualizers allow timbre to be much more analyzable than it was in the past. This dissertation takes advantage of these technological developments in order to illuminate the musical innovation represented in today's commercial pop music.

CHAPTER 3

(DYS)FUNCTIONAL HARMONY: HANDING OVER THE ROLE OF HARMONIC FUNCTION TO SOUND PRODUCTION

When I play a song to someone and ask ‘So how do you like this?’, I don’t care all that much about what they say. What I really pay attention to is how they act, their body language [while they listen to the song]. People who lose their concentration give themselves away very quickly. If they start fiddling with their phones as the second verse kicks in, there may be something about the tune that wasn’t good enough. Something also happens when I listen as if with other people’s ears. I get nervous and think to myself, ‘Shit, this part is a bit too slow’.⁸

Max Martin, 2016

Max Martin, the “Obi-Wan of pop songcraft,” according to John Seabrook,⁹ wants you to pay attention. He is not satisfied if you only casually like his song. He wants you to be captivated from start to finish. If at any point you start to lose focus it means that something is wrong with the song and it must be fixed. In an age when each pop song competes for the attention of the masses – not only with hundreds of other songs that are released every year, but also with

⁸ Jan Gradvall, “World Exclusive: Max Martin, #1 Hitmaker.” *Dagens industri Långläsning*. 2016.

⁹ John Seabrook, “The Doctor is In: A Technique for Producing No. 1 Songs.” *The New Yorker*. 2013.

numerous other forms of digital-age distractions – Martin’s approach makes sense. But this is easier said than done. How can a songwriter make sure that millions of people whose cultural backgrounds and predispositions are diverse and unknown are all enchanted by the same song? Of course, even Martin would not expect every person on Earth to fall in love with his songs, but if we take the view count of YouTube videos as an indicator of popularity, his songs do have significant mass appeal, as evidenced by the hundreds of millions of views garnered by many of his big hits.

Is there a formula for appealing to the masses with pop songs in the way that many hits co-written and produced by Martin, such as Katy Perry’s “Teenage Dream” (2010) and Taylor Swift’s “Blank Space” (2014), have done? According to him, such a formula does not exist. However, he does concede that there is a toolbox from which he draws when working on a song. One of the tools Martin mentions is creating a sense of familiarity. From the same interview by Jan Gradvall:

For instance, take ‘I Wanna Be Your Lover’ with Prince. The verse and chorus of that song are exactly the same. But as a listener, you don’t really notice since the energy of the chorus is completely different compared to the verse... Once the chorus comes, you feel like you’ve heard it before. And you have! You’ve heard it in the verse. It automatically creates a sense of familiarity. Prince does this a lot. ‘Let’s Go Crazy,’ same thing. I’ve used this trick a few times myself. In ‘Do You Know (What It Takes)’ with Robyn for instance.

It is not hard to imagine why familiarity is an effective tool for keeping listeners’ attention. It helps us to sing along and feel engaged, becoming active participants in the musical experience. Part of this feeling of engagement is a result of our ability to anticipate forthcoming musical events at almost any given moment. Most of today’s hit songs try to make sure that we are in a constant state of anticipation, whether through familiarity created by repetition or via

other means that guide the listener through a song. This is not new. When the music moves toward a perfect authentic cadence in a common-practice tonal piece, the dominant triad or seventh chord causes us to anticipate an arrival at a tonic resolution. When thematic material is recycled, our familiarity with it helps us to predict important moments before they occur. However, while the instances in which we feel a heightened sense of anticipation toward the inevitable are reserved for key areas in a traditionally tonal piece, pop hits aim to keep us on relatively high alert over the course of an entire song. Of course, there are moments in every pop song in which the sense of anticipation is at its peak compared to the rest of the song, but as I will show in this chapter and in Chapter 4, each section of a typical pop song has anticipatory properties that help to keep the listener engaged throughout. Moreover, anticipatory gestures in contemporary pop music, such as filter sweeps, drum intensification, and abrupt removal of sonic layers, more easily propel the listener toward a goal (compared to tonal gestures) due to their explicitly directional nature. Traditional tonal functions, by contrast, become more effective as we gain experience as listeners and become familiar with the patterns from which they emerge.

This dissertation predominantly explores the emergence of sound production as a primary tool for creating syntax based on the idea of tension and release in contemporary pop music. In this chapter, I discuss the effect of this development on harmony and pitch material in this genre. My core arguments are that:

- a) Contemporary pop music has largely abandoned harmonic function as a means of achieving tension and release, as a result of relying on non-pitch-based sound production techniques that perform similar functions. Chord changes in this music act as agents of harmonic color and as temporal markers, but a V, for example, is not a dominant in the anticipatory sense.

- b) Tonal tendencies are therefore generally avoided in today's pop.
- c) While there are many similarities between the functions performed by sound production techniques and traditional tonal functions, current pop songs do not generally achieve closure similar to the closing tonic.

I will demonstrate these points by discussing the components that go into creating anticipation (which correlates with increasing tension), in part relying on David Huron's observations from his book, *Sweet Anticipation* (2006), and comparing how these components manifest themselves in different genres. As part of this discussion, I will explore the ways in which pop music avoids activating these components in its pitch matter, and instead uses a range of non-tonal gestures to create a sense of anticipation.¹⁰

THE EVOLUTION OF POP MUSIC

At first glance, contemporary pop music seems to have broken away from the evolutionary line of common-practice tonal music. While pop songs still have tonal centers and use diatonic pitch collections as the foundation for the foreground material (such as the vocals and instrumental or synthesized accompaniment), many of them have done away with core tonal idioms that cause the listener to feel anticipation, such as cadential progressions, chromatic alterations, and "activated" tendency tones. Nevertheless, I will show in this chapter (as well as in Chapter 4) that mechanisms borrowed from electronic dance music (EDM) fulfill anticipatory functions similar to those in common-practice tonal music. In order to do this, I will draw comparisons between common-practice tonal music, 20th-century rock/pop music, and

¹⁰ The term "non-tonal gestures" refers to gestures that do not rely on pitch, such as filter sweeps and drum intensification, as well as pitched gestures that are not functionally tonal, such as pitch bends (*glissandi*).

contemporary pop music, and explore how each of these genres deals with the idea of tension and release. For the sake of clarity and due to limitations of scope, I generalize the first two and exclude some of their nuances and complexities. For instance, I make assertions regarding common-practice tonality that fit a textbook description of this music, while acknowledging that this representation does not capture the nuance that is present in many individual pieces. I also make the generalization that cadential closure is inherent in 20th-century rock and pop music, though examples to the contrary surely exist. My purpose is not to caricaturize these musical genres, but to show how the reduced ubiquity of certain tonal features in contemporary pop music compared to its tonal predecessors has allowed sound production to assume a primary role in generating tension and release.

Terminology

The concepts of *tension* and *anticipation* are central to this dissertation. These terms are mostly interchangeable within the framework of this topic, but I sometimes choose one rather than the other, depending on context. David Huron equates tension with “a slight amount of psychological stress.”¹¹ This psychological stress correlates with a conscious or semi-conscious feeling of anticipation. In this sense, *tension* implies a visceral reaction, while *anticipation* suggests an intellectual experience, but the correlation between the two is so high that one can be substituted for the other. However, *tension* is also idiomatically used to describe a musical trait. For example, the moment prior to the resolution of a dissonant sonority is considered a tense one; adding rhythmic layers is often perceived as increasing the tension; a filter sweep gradually increases tension. So, in this context, when referring to tension as a musical property, it is not interchangeable with anticipation (although there is a high level of correlation between the two).

¹¹ David Huron, *Sweet Anticipation* [Cambridge: Bradford Books, 2006], 307.

In addition, I frequently use the terms *sonic density* and *sonic energy* to discuss texture and intensity. While there is a high degree of correlation between these two terms as well, they are also not synonymous. I use the term *sonic density* when evaluating the presence (or loudness) of frequencies across the sonic spectrum. For example, a thin texture that includes only sparse piano chords and a single vocal line is relatively low in sonic density, whereas a texture that includes drums, bass, synth pads, various background sounds, and multiple vocal tracks is sonically very dense. Some of the techniques that I lay out in this dissertation affect tension by altering sonic density. For instance, a filter sweep normally increases sonic density by releasing or amplifying frequencies that were previously cut or muffled.

The term *sonic energy* refers to the total intensity of the sonic activity, including sonic density, rhythmic intensity, and pitch register. It would be difficult to discuss sonic energy in absolute terms, because the various elements that amount to sonic energy are measured by different parameters (for instance, rhythmic intensity is measured by the rate of attacks, while sonic density is measured by the loudness of frequencies across the sonic spectrum). However, for the purpose of analyzing a song, it is usually sufficient to note the comparative level of sonic energy between musical sections, as well as gradual changes that occur within a section. Increasing the sonic density contributes to elevating sonic energy, as does increasing rhythmic intensity, and in some contexts, extending the vocal range. In this chapter, I generally use the term “sonic energy” more frequently than “sonic density” because more often than not, changes in sonic density coincide with changes in rhythmic intensity and pitch register, necessitating a reference to the totality of sonic activity rather than to just the sonic density. Moreover, as I mentioned earlier, increasing sonic density usually correlates with increasing tension, but the two

are not synonymous. Abrupt drops in sonic energy, for instance, usually increase tension rather than decrease it.

Non-Anticipatory Harmony

Nicki Minaj’s “The Night is Still Young” (2014) features a recurring chord loop consisting of the triads A, C#m, B, and F#m (IV-vi-V-ii in E major), with syncopated connective chords played briefly just prior to the downbeat of the next triad. This song is clearly in E-major, as indicated by the collection of triads that make up the chord loop, as well as by the vocal melody of the chorus (see Example 3.1). However, there is no tonic triad heard on a strong beat throughout the entire song, which is highly unusual. The E-major triad appears only as a connecting chord on a weak beat. In fact, had the connecting chords been entirely omitted, while the song would lose some of its musical nuance, this omission would not affect its perceived tonality, nor would it alter any harmonic function.

Vocals (starting at 0:45)

The night is still young The night is still young The

Accompanying harmony IV (V) vi (I)

night is still young and So are we

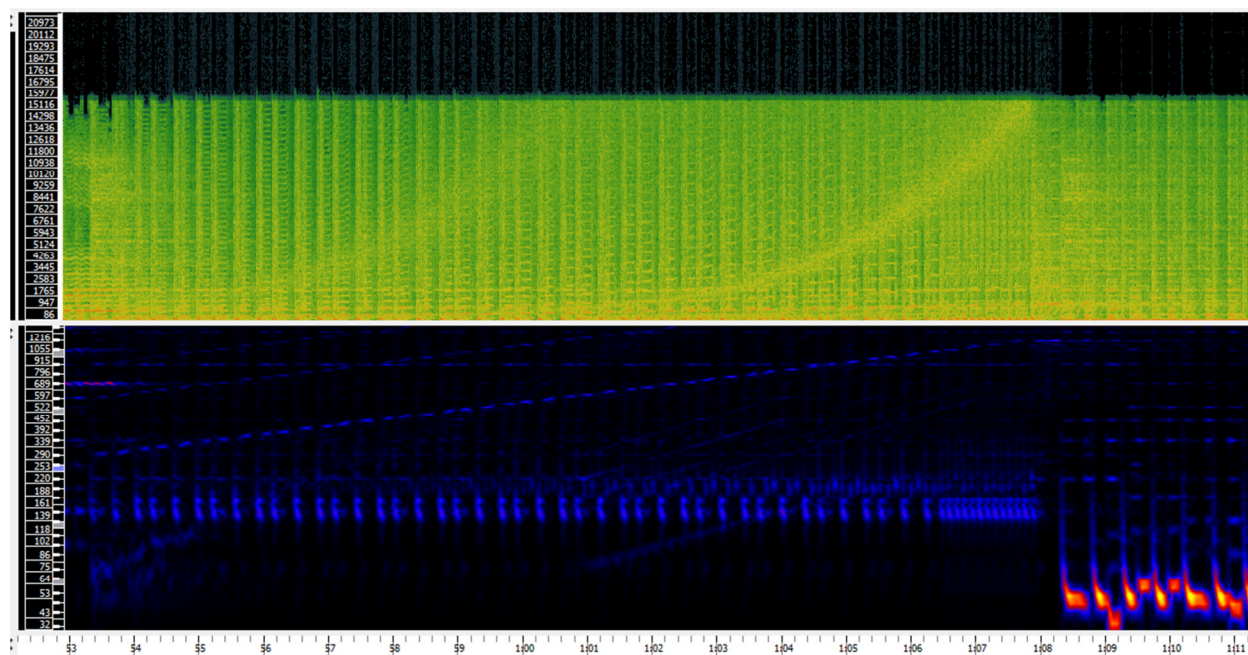
V (I) ii (I6)

Example 3.1. Nicki Minaj, “The Night is Still Young” (2014). Vocal melody and accompanying triads (represented by Roman numerals) in the first part of the chorus (0:45-1:00).

Though most chord loops in today's pop music include the tonic triad, its absence in this song can easily go unnoticed because it is congruent with a salient trait of this genre – it generally avoids complete closure. This may seem counterintuitive because goal-oriented risers (intense dance-oriented buildups) resolving at the *drop* are staples of today's pop, due to the influence of EDM. The drop is the moment at which high-amplitude low-frequency sounds (e.g., kick drum and bass) are abruptly reintroduced into the relative foreground of the mix after temporarily being removed or obscured. The riser leading up to the drop usually includes a filter sweep, which gradually brings higher and higher frequencies to the foreground, highlighting a contrast with the imminent drop. To a certain extent, the arrival of the drop feels like a cadence, in the sense that it provides a highly anticipatory passage with the predicted resolution. However, the drop is not an ending, but rather a beginning of a highly charged passage, and does not provide the type of absolute closure associated with the traditional function of the closing tonic.

The spectrogram in Example 3.2 visualizes important components of a riser-drop sequence from Rihanna's "We Found Love" (2011). This is a two-part riser, although some elements are continuous through both parts. The plain spectrogram shows two ascending yellow arcs, at 0:53-1:00 and 1:00-1:08, which reflect the noise sweeps (filter sweeps applied to white noise) participating in the buildup of sonic energy. Their trajectory is identical, but the second arc is more pronounced, meaning that the same ascending frequencies are now louder, as the drop is approached. The melodic-range spectrogram, which zooms in on the range of fundamental tones, shows a clearer picture of several sounds. The tom and snare activity at ~130-230 Hz continue throughout the buildup, intensifying to sixteenth notes leading up to the drop. A continuous rhythmic horn sound is represented by an ascending broken line, starting at ~300 Hz and peaking at ~1100 Hz. To supplement this line, a continuous ascending pitch bend,

starting at ~75 Hz with independently-audible harmonics, is introduced at the start of the second half of the buildup (1:00). The components of the buildup process – the second noise sweep being louder than the first, the rhythmic horn reaching higher frequencies, the drum intensification towards the end, and the adding of the pitch bend in the second half – all contribute to the growing sense of anticipation by guiding the listener toward the drop. The meter and hypermeter are crucial to this process, since they establish the temporal patterns that allow the listener to confidently predict the precise timing of the drop. The tension peaks at the break (an abrupt pause of nearly all sonic activity), which occurs just prior to the drop, which finally arrives at 1:08. The drop resolves the tension that had been built up in the preceding 15 seconds, but unlike an arrival at a closing tonic, the music further intensifies by sustaining a high level of sonic energy rather than relaxing, bringing the tension to an even higher level. This burst of musical energy is the most common result of such a buildup, and contrasts with traditional tonal models, in which a tonic sonority that resolves tension built up by a dominant sonority signals a relaxation of the music, rather than intensification.



Example 3.2. Rihanna, “We Found Love” (2011). Spectrogram visualization of riser (0:53-1:08).

Still, although their musical behavior and mechanisms for achieving tension and release are vastly different, there are similarities between the sonic syntax expressed in the riser-drop sequence in “We Found Love” and the tonal syntax of a perfect authentic cadence in a piece by Mozart or Haydn. Both induce tension based on the listener’s ability to anticipate a resolution at a particular moment. The cadence’s resolution is the tonic, while the riser’s resolution is the drop.

There is also an analogy to be made regarding the varying degrees of intensity to which these tension-and-release dynamics occur in both traditional tonal music and contemporary pop music. A dominant-tonic resolution can be heard at the end of local phrases, as well as at important structural moments, such as when a dominant expansion arrives at the recapitulation in a sonata form movement. Likewise, a small noise sweep at the end of a verse can lead to the

arrival of the pre-chorus, while large-scale sweeps such as the one discussed above usually occur at the most climactic moment in a song. Although all of these scenarios represent increasing tension, the level of intensity varies due to the differences in magnitude.

THE COMPONENTS OF ANTICIPATION

David Huron suggests that predicting a forthcoming musical event with high probability of success will cause us to feel strong anticipation toward its arrival. This anticipation builds up tension, in the form of slight psychological stress, as the event approaches. Huron asserts:

“In light of the brain’s disposition toward statistical learning, it is likely that the feeling of anticipation is greatest when the probability of an event approaches certainty. For zeroth-order probabilities, it is not common for events to be certain. Without some context, there are usually many possible events. However, many first-order probabilities really do approach statistical certainty. For Western-enculturated listeners, examples of strong feelings of anticipation include the feeling that a chromatic tone should resolve to a diatonic neighbor, and the feeling that a drum fill should lead to a hypermetric downbeat.”¹²

Huron emphasizes the role of certainty in creating a feeling of anticipation. The more confident we feel about our prediction of the arrival of a musical event, the greater the tension and feeling of anticipation. In other words, the more we can narrow the options for the possible character and timing of the impending musical event, the more anticipation we will feel toward its arrival. Although we cannot quantify the level of anticipation, we can assume that if there are infinite options, the level of anticipation would be zero, and if there is a single inevitable option, the level of anticipation would be at its maximum. Between these extremes, there is a spectrum

¹² Huron, *Sweet Anticipation*, 306.

in which numerous options would yield a low level of anticipation (but not zero), and as we limit the options, the level of anticipation becomes higher.

The role of context in achieving a high level of certainty is also stressed by Huron. Context is crucial in determining the two key factors in anticipating a forthcoming event – the *what* and the *when*. Among the elements that make up the musical context are a) general musical surroundings; b) stylistic norms and idioms; c) the internal makeup of a specific piece; and d) meter. I will briefly discuss how each of these factors contributes to creating a sense of anticipation, with the caveat that none of these categories are isolated. General musical surroundings would mean little without established stylistic idioms or piece-specific themes that endow them with their meaning. Since music is temporal, meter is a crucial factor in establishing almost every stylistic norm and thematic idea in the genres discussed in this dissertation. Nonetheless, it is worth elaborating on each of these categories separately, in order to better understand the process of generating a sense of anticipation, and how this translates from one genre to another.

General Musical Surroundings

Imagine an out-of-context root position D-major triad, such as the one in Example 3.3.a. Such a sonority played out of the blue would not generate a significant level of anticipation because we do not have sufficient context to predict what we are going to hear next and when we will hear it. In Example 3.3.b, we introduce several root position triads at regular time intervals prior to the D-major triad. Due to the regular time intervals, we are likely to anticipate an additional chord exactly two seconds after the D-major triad has sounded. As for its character, we will expect this chord to be a root-position triad, because that has been a trait common to all

of the chords in the series. However, we will not have a particular expectation as to whether it will be a major or minor triad, and what its root will be, because a reliable pattern has yet to be established in those respects. By contrast, in Example 3.3.c, we will anticipate that the next sound will be an E \flat -major triad, due to the established pattern of root-position major triads ascending by half-steps. We will not, however, feel confident in predicting the timing of its attack, since the triads preceding the D-major triad sounded at seemingly random time intervals within a one-minute time frame. In Example 3.3.d, we will feel the strongest sense of anticipation toward the sonority that follows the D-major triad. Since a pattern has been established for both the *what* and the *when*, we are likely to anticipate that an E \flat -major triad will be played exactly two seconds following the D-major triad.

a) Out of context.



b) Pattern establishing timing and general, but not specific, character of next sonority.



Example 3.3. *Patterns that contribute to anticipation.*

♩ = 120

B \flat major, E \flat major, F \sharp m, B \flat major, F \sharp m, ?

Stylistic Norms and Idioms

50

tonic, because it conforms to our prior experiences of listening to other Mozart pieces. We are confident that these outcomes are forthcoming, and we feel emotionally rewarded when our prediction comes true.

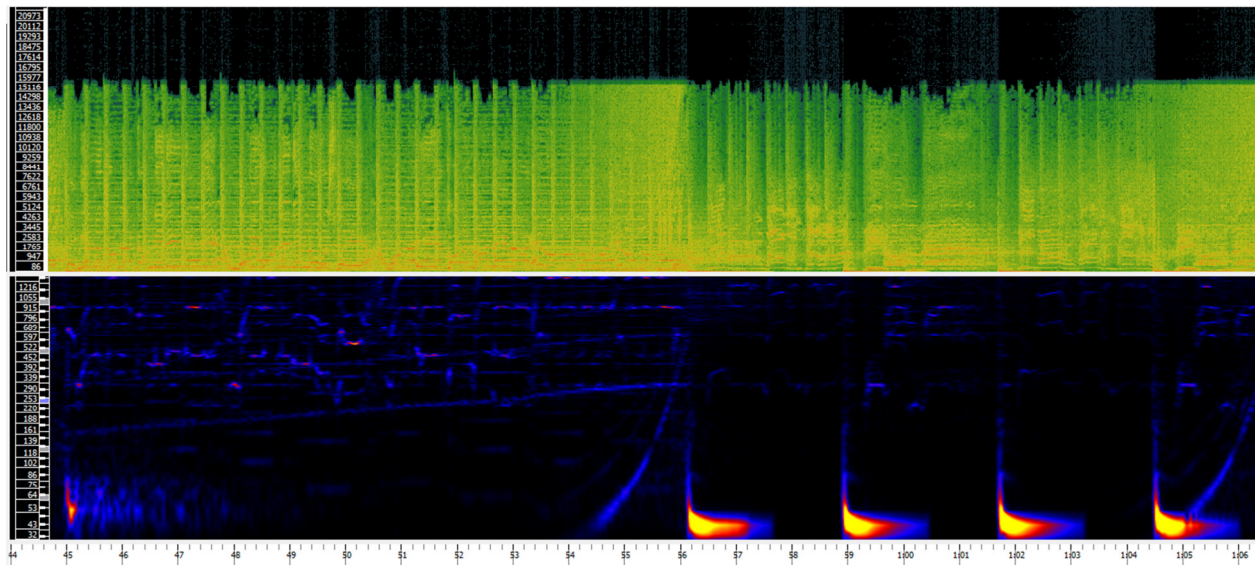
Once a piece has established a set of traits that fit within a certain genre we expect that it will continue to behave accordingly. These traits may include timbre, musical objects (e.g., triads, tone clusters, power chords), and interactions between these musical objects. The timbre of distorted electric guitars, combined with certain drum and bass patterns, may cause us to expect that the music will behave like a rock song, which means that a guitar solo after the second chorus will be a viable option for us to anticipate, for instance. If we are listening to a Mozart piece, we would not think of a syncopated cadential resolution arriving on a weak beat as a realistic possibility, but it would certainly be an option if we were listening to a Duke Ellington big-band arrangement. The more familiar we are with a genre and the more a piece adheres to the “rules” of the genre, the easier it becomes to elicit expectations.

Internal Makeup of a Piece

Nearly every piece of music establishes themes and behaviors that we learn to anticipate as the music progresses. Example 3.4 shows spectrogram excerpts of three buildup-drop sequences in David Guetta’s “Hey Mama” (2014, featuring Nicki Minaj and Bebe Rexha). While most buildups in this genre that include ascending sweeping gestures are followed by a climactic, sonically dense section, this song establishes a different pattern. The end of the first chorus (0:44-0:56) features a riser that steadily increases tension by using a continuous ascending pitch bend, which is joined by an additional, lower pitch bend and a noise sweep toward its end. Drawing on stylistic idioms, we would expect a drop that triggers a highly energetic section.

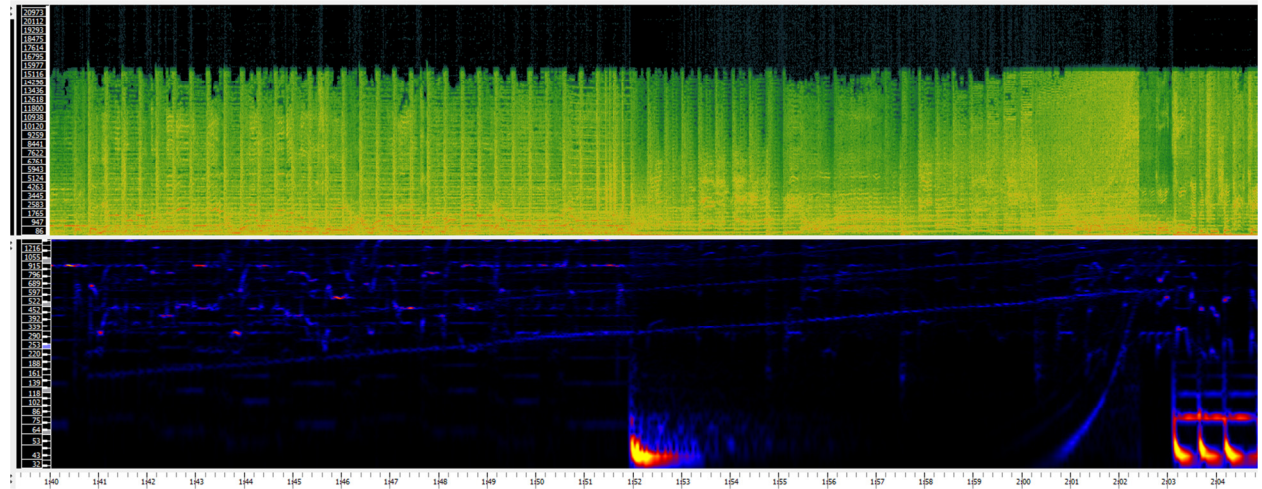
However, while the anticipated drop indeed arrives, marked by an emphatic bass-drum hit, the section it triggers (0:56-1:07) is surprisingly hollow, and acts more as a transition than a moment of arrival. As the end of the second chorus is approached with a similar pitch bend (1:40-1:52), such a “deceptive” result becomes a much more likely option in our bank of viable possibilities. And indeed, not only is the following section similarly hollow, but the pitch bend also continues to ascend, increasing the tension even further toward what ends up being an additional drop upon arriving at Nicki Minaj’s rap verse (2:03). By the end of the third chorus, as the same pitch bend ascends (2:48-2:59), we already expect a similar, sonically hollow section with a relatively high degree of certainty, and indeed, the following section confirms this expectation with a section identical to the one we heard in 0:56-1:07.

a) 0:45-1:06

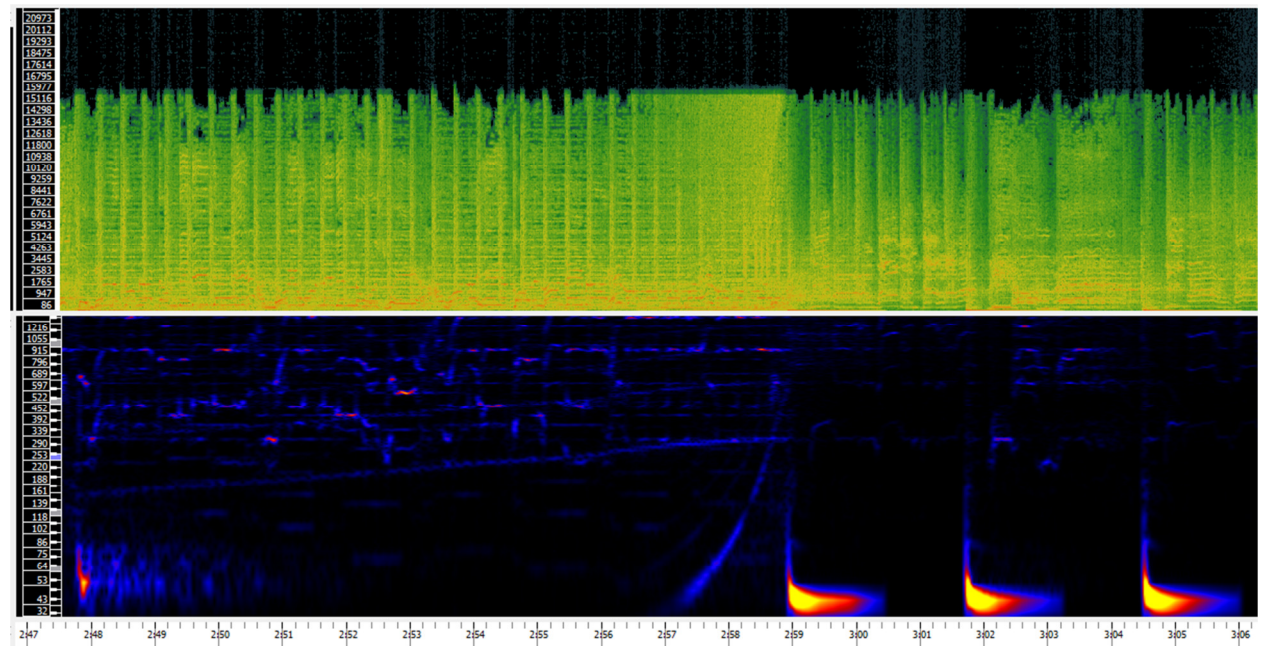


Example 3.4. David Guetta, “Hey Mama” (2014), featuring Nicki Minaj and Bebe Rexha. ‘Deceptive’ drops.

b) 1:40-2:05



c) 2:48-3:06



Example 3.4. *Cont.*

Of course, the options being planted in our minds are not limited to exact repetition. Music tends to give us hints regarding what we can expect. The examples above are part of a sonic theme that defines the character of “Hey Mama.” For instance, the intro (0:00-0:11) also builds up anticipation toward a drop via an ascending noise sweep. The verse indeed starts with a drop, but the only participant in the drop is the bass, with relatively high-pitched percussion and an otherwise thin texture accompanying the vocals. This is not identical to the recurring hollow sequences between the various iterations of the chorus and the following refrains, but their characteristics are similar. The intro-verse sequence sets up a sonic theme that ties together interactions that are similar to one another, albeit not identical.

Meter

Meter is perhaps the most crucial element in generating a sense of anticipation, because it governs the timing of the arrival of the anticipated event. Huron suggests that our feeling of anticipation increases the closer we move to the precise time at which we anticipate the arrival of an event. If it arrives too early, “the tension response will fail to reach its potential peak” (2006, 314). If it fails to arrive at the anticipated moment, the tension will drop until another possible moment of arrival approaches. Therefore, in order for the listener to achieve the most fulfilling tension/release experience, s/he must feel confident in precisely predicting the possible points of arrival.

Meter is responsible for helping us determine and anticipate these time points. The resolution of a structural cadence or a sonic buildup is typically expected on a metric/hypermetric downbeat. While a delayed arrival can ultimately result in a more satisfying resolution, it still has to occur on a perceived downbeat. A classical performer may slow down

prior to a cadence in order to emphasize it and increase the feeling of anticipation, but the arrival will still be perceived as occurring on the downbeat. A contemporary pop song will not typically slow down for a similar effect, but rather will extend the buildup by a precise measure of time that will ensure that the arrival still occurs on a downbeat.

Additional Factors

The above parameters contribute to a feeling of certainty not only by limiting the number of viable options, but also by establishing a hierarchy between them. For instance, in common-practice tonal music, V is so frequently followed by I that in most circumstances the sound of V causes us to automatically anticipate I, even though other options (e.g., vi) are perfectly acceptable within the boundaries of the genre. Therefore, even though a number of viable options exist, I is so much higher in the hierarchy of possibilities, that we anticipate it by default. We refer to other options as “deceptive” or “evasive,” even when we actually anticipate them (such as in a repetition of a passage). By contrast, when there are several equally viable options, such as when moving from an initial tonic, our sense of anticipation is diminished in comparison. Perhaps the strongest sense of anticipation occurs at important structural moments, because these moments are often defined by a very specific outcome, and therefore this outcome is at the top of the hierarchy of options by a wide margin. For example, key moments in classical form are characterized by a perfect authentic cadence. Therefore, once this cadence begins to unfold, our feeling of anticipation is at a peak because we can predict a resolution to the tonic with a great degree of certainty.

It is worth emphasizing that since the feeling of anticipation occurs during the phase in which we predict the outcome, whether or not our prediction proves to be correct is immaterial to

the level of anticipation (although it gains significance when we hear the same or similar music repeated). An anticipated event might be replaced by a completely different one at an unforeseen time, but this has no bearing on our prior expectations. For instance, the laws of physics do not prevent a riser such as the one in “We Found Love” from being prematurely interrupted by a jazz piano solo. However, this is so improbable that it does not register as a viable option in our mind as the riser unfolds, and does not affect our anticipation of the drop until the interruption occurs. The patterns we hear in the music prior to the present moment (both in the present piece and in our history as listeners) help us to arrive at a limited number of viable possible options toward which we can feel a sense of anticipation, obscuring the infinite number of highly unlikely possibilities.

ANTICIPATORY FUNCTIONS

Harmonic functions, beyond their syntactical role, represent varying degrees of anticipation in most tonal music. For example, a dominant function is more anticipatory than a tonic function, due to the expectation it generates for a specific outcome, as described above. Moreover, a single function can vary in its anticipatory properties in different contexts. Similar variations exist in some of the syntactical features of contemporary pop music.

I group sonic functions into three major categories – the *setup*, the *buildup*, and the *peak*. All three of these categories are anticipatory to some degree, and can be further divided into subcategories. For instance, a riser such as the one in “We Found Love” is a highly anticipatory buildup because we can confidently predict the arrival of the drop at the hypermetric downbeat. Conversely, adding a hi-hat to an existing texture also constitutes a buildup, but a much less

directed one, since it is not as explicitly goal-oriented and the hierarchy of options for a following musical event is not nearly as clear as in the case of a riser-drop sequence. Even similar or identical gestures can vary in intensity, depending on context and magnitude.

I explore and expand on these sonic functions and their internal variations in chapters 4 and 5. However, my purpose in introducing them in this chapter is to show the syntactical similarities between tonal harmonic functions and pop's sonic functions, and to illustrate the shift from the former to the latter as chief dramatic drivers in this music. Since pop music is still a tonal genre, despite the emerging dominance of the sonic syntax, the significance of these similarities is expressed not only in the creation of the new sonic language, but also in the altering of the tonal language. Sound production has "freed" harmony from the need to perform its traditional functions, and therefore allowed the harmonic syntax in this music to change dramatically while retaining its basic properties, such as diatonic pitch collections and major/minor triads.

Harmony in 20th-Century Rock and Pop Music

Twentieth-century rock and pop models have fiddled with the notion of what constitutes a dominant or predominant, but they still largely operate on the common-practice principle, wherein harmony strives to return to the tonic. For example, in George Michael's "Freedom 90" (1990), after the intro establishes C major as the tonic with a C-B \flat -F-C, or I- \flat VII-IV-I chord loop (which reappears as the chorus chord loop), the verse begins on a G-major triad (0:52), which may cause us to momentarily feel like we have shifted to a different key area. But, while the new chord loop resembles a version of the loop we heard in the intro transposed down by a perfect fourth, beginning with G-F-C (or V-IV-I), it remains on the C-major triad rather than loop back to G, Making C-major the closing triad of both loops. Moreover, the melodic phrases throughout

the song are structured to end upon the arrival of the C-major triad. At key moments, the cadential feeling is emphasized by the melody ending on C, such as at the end of the verse (see Example 3.5.a) and at the end of the first part of the chorus (see Example 3.5.b). The chorus ends with the lead vocal on C as well, although at that point we hear a choir in the foreground, obscuring the lead vocal with G in the soprano voice (Example 3.5.c). Nevertheless, the closure achieved by moving to the C-major triad and the tonic note in the melody throughout this song (in addition to the C-minor vamp in the pre-chorus) shows that while this song defines the dominant and predominant functions differently than in a traditional tonal piece, the principles of harmonic functionality are still primary tools for achieving tension (and as a result, a feeling of anticipation), which is resolved at the tonic.

a) End of verse I (1:23-1:34).

But to - day the way I play the game is not the same — no way Think I'm gon-na get me some hap-py

Example 3.5. *George Michael, “Freedom 90” (1990).*

b) Chorus I – part 1 (1:55-2:16).

C B \flat F C

All we have to do now is take these lies and make them true some - how

C B \flat F C

All we have to see is that I don't be-long to you and you don't be-long to me Yeah Yeah

c) Chorus I – part 2 (2:16-2:38).

C B \flat

Choir Free - dom Free - dom

Lead vocal got. I won't let you down so please don't give me up

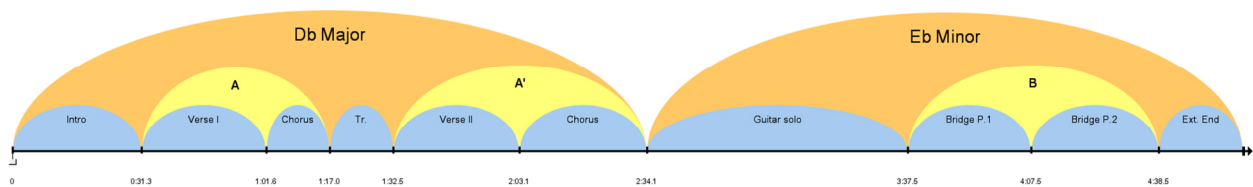
F C

Free - dom You got to give for what you take

'cause I would real - ly real - ly love to stick a-round

Example 3.5. (cont.).

Another example of cadential closure in rock can be heard in Guns N' Roses's "Sweet Child O' Mine" (1987). The first half of this song (0:00-2:33) is in D \flat -major and consists of the verse-chorus pairs. The second half (2:34-4:59) is in E \flat -minor and includes a long guitar solo and an extended bridge that does not return to the chorus (see Example 3.6 for form timeline). In both parts, important moments end in full cadential closure. The chorus (1:01-1:17) is made up of two vocal phrases, each ending upon the arrival of the tonic triad (See Example 3.7). The first phrase concludes on $\hat{3}$, setting up the second phrase to achieve full closure on $\hat{1}$. The bridge ends with a series of power chords (perfect fifths played in parallel motion on a distorted electric guitar) that climb up the E \flat -minor pentatonic scale to achieve closure on E \flat (4:30-4:38). This note is then held for an additional eight seconds by the lead vocalist, Axl Rose, who momentarily dips it down to D \flat and back up to E \flat , to emphasize the closure.



Example 3.6. *Guns N' Roses, "Sweet Child O' Mine" (1987).*

Example 3.7. *Guns N' Roses, "Sweet Child O' Mine." Transcription of first chorus (1:01-1:17).*

Twentieth-century popular music, as demonstrated by these songs and many others, challenges the notion that a V chord, and particularly the leading tone, is necessary for achieving closure. However, while it is not expressed via perfect authentic cadences, the idea of cadential closure as a way of delineating form and emphasizing endings remains intact. All of the chord loops in both "Freedom 90" and "Sweet Child O' Mine" end on I. The vast majority of chord loops in these genres either begin or end on I, in a way that facilitates cadential closure. Moreover, nearly every song in these genres ends on the tonic, unless the outro fades out. The chord loops featured in the two aforementioned songs – I- \flat VII-IV-I, V-IV-I-(I), and V- \flat VII-I – are quite common in rock music. By contrast, chord loops in recent pop music include the following:

- i-III-VI – Britney Spears, "Till the World Ends" (2011).
- i-III-VI-iv – Ludacris, "Representin'," featuring Kelly Rowland (2012)
- I-ii-vi – Daya, "Hide Away" (2015).

- VI-iv-I-VII – Demi Lovato, “Cool for the Summer” (2015) & Justin Bieber, “Beauty and a Beat” (2012).
- I-vi-IV-vi – Jason Derulo, “Want to Want Me” (2015)
- IV-vi-V – Justin Bieber, “Sorry” (2015)
- IV-vi-V-ii – Nicki Minaj, “The Night is Still Young” (2015)
- I-IV-vi – Miley Cyrus, “Adore You” (2013)

The contrast of character between the above chord loops and those common in 20th-century rock and pop are indicative of core differences between the genres. Although both genres rely largely on repeating chord loops framed in four-measure blocks, it is not a coincidence that the contents of the loops differ to a great degree. While chord loops common in rock music are constructed in a way that facilitates cadential closure, chord loops in pop seem to strip the tonic triad of its closural properties. Some of these loops, like the ones in “Sorry” and “The Night is Still Young,” avoid the tonic triad altogether. Others start with the tonic triad but avoid V or any other chord that could act as a dominant, or a cadence-oriented pre-tonic, such as a \flat VII or IV, as the fourth chord of the loop. This lack of cadential closure is a central characteristic of today’s pop.

Harmony in Contemporary Pop Music

Revisiting Example 3.1, the chorus of “The Night is Still Young” does achieve melodic closure, ending on $\hat{1}$. In fact, the vocal melody of the chorus consists almost entirely of scale degrees belonging to the tonic triad. However, the underlying harmonic accompaniment, IV-vi-V-ii, does not “agree” with the melody, nor does it allow it to achieve the types of cadential arrival common in traditional tonality and rock music, which require the tonic triad to underscore

the resolution. The disconnect between the melody and underlying chord loop is not new, and has been addressed by Allan Moore (1995, 189), who termed it a “‘divorce’ between melodic and surface harmonic schemes which we might characterize as ‘historically mediated slippage’ ...” Moore explains this divorce as having roots in earlier blues, in which each note of the pentatonic scale is “compatible” with the underlying accompaniment. Drew Nobile (2015) further unpacks the idea of a “divorce” between melody and harmony by suggesting three types of such divorce, one of which is the “loop divorce,” wherein the chord loop does not represent goal-oriented harmonic motion. Nobile argues that in this case, the melody can create independent structures that achieve closure.

While I concur with Nobile’s assertion that when viewed as a separate entity, melodies can achieve closure independently of the underlying harmony, I argue that this type of closure is not comparable in its completeness to cadences in which the harmony and melody are in sync. However, it is important to note that although he provides a few examples of fairly recent pop music, Nobile focuses primarily on 20th-century rock and pop, which partially explains where he and I differ in our interpretation of this phenomenon. I agree that the final note in the chorus of “The Night is Still Young” ($\hat{1}$) feels relatively stable, certainly more stable than we would normally expect $\hat{1}$ to sound over a ii triad. However, I do not hear it as representing a type of closure similar to that of a perfect authentic cadence, in which all of the primary properties cooperate to maximize the sense of closure. Nor do I hear it as similarly stable to the cadences in “Freedom 90” or “Sweet Child O’ Mine.” Moreover, Nobile argues that “This conflict usually ends when the harmony escapes its loop and joins the melody’s structural motion.” I find this not to be the case in most contemporary pop songs. The most typical way in which most current pop songs end is by abruptly stopping the music after the final chorus or post-chorus, leaving only an

echo or a soft delayed repetition momentarily trailing. There is no attempt at closure, perhaps due to a desire to keep the listener ready for the next song. These differences reflect a shift in the spirit of the genre. While 20th-century rock and pop inherited the concept of closure as a tonal goal and a means of delineating formal sections, contemporary pop retains the idea of creating anticipation toward a goal, but forgoes full closure.

I should clarify – the chord loops common in contemporary pop music do not create the state of perpetual anticipation. Rather, they foster it by not conforming to the traditional role of harmony as directional and goal-oriented. On the surface, the makeup of the typical chord loops in today’s pop music may not seem revolutionary: repeated loops of three or four chords that accompany the vocal melody are abundant in 20th-century rock and pop; individually, the chords used in current pop songs are quite ordinary as well; and in fact, in public discussions of pop music, such as online forums, it is a common complaint that harmony in pop music is a “dumbed down” version of harmony in older popular music. However, while it is true that many of the harmonic complexities common in the music of the Beatles or Queen are extremely rare in today’s pop, I argue that the things that harmony does not do in modern pop reflect deliberate decisions and an interesting shift in the thought process of creators of this music.

The significance of what harmony does not do in contemporary pop music requires some unpacking. I have discussed its general avoidance of structural cadences and goal oriented functionality. However, there are additional features of tonal harmony common in 20th-century rock and pop that are rarely heard in today’s pop music. These include:

- Dominant seventh chords
- Chromatic alterations

- Harmonic/melodic minor alterations
- Meaningful (functional) chord loop changes between sections

The first three items on this list represent features of traditional harmony that are responsible for generating a feeling of anticipation toward a predictable resolution. The fourth item represents the role of harmony in form delineation. In rock songs, the progressions in the verse and the chorus (as well as the pre-chorus, if the song contains one) usually differ, and at times these progressions represent functions in a large-scale harmonic structure, as Nobile explains.¹³ Today's pop songs tend to keep the same chord loop throughout and allow the lyrics, melody, and sonic texture to create the form. Even in songs in which there are changes in the chord loop between sections, these changes are not functional and do not constitute a large-scale harmonic arc.

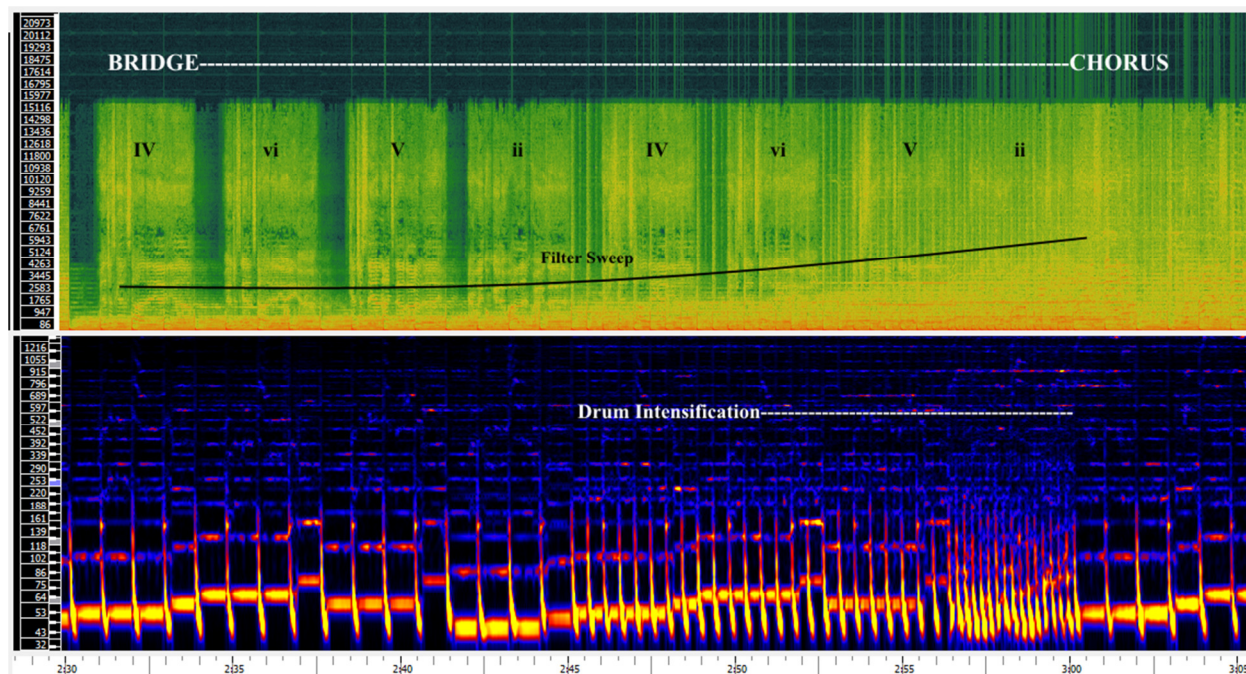
In other words, while the chords in a typical chord loop indeed represent a pitch collection with a tonal center, they tend to avoid performing their traditional functions, such as dictating tendencies or defining the form of a song. Two things happen as a result of the “emancipation” of harmony from its functional roles: a) other musical aspects, such as timbre, gesture, and spacialization, in conjunction with meter, assume the anticipatory roles, as I have shown earlier in this chapter; and b) new chord combinations that are not directed toward a resolution are introduced to the language of pop music.

The bridge (2:30-3:00) in “The Night is Still Young” exemplifies the results of harmony's changed role. The chord loop IV-vi-V-ii, or A-C#m-B-F#m, does not possess any anticipatory properties. The third chord of the loop, V, does not generate its traditional

¹³ Drew Nobile, “Verse-Chorus Forms as Harmonic Patterns” (Paper presented at the annual meeting for the Society for Music Theory, St. Louis, Missouri, October 29-November 1, 2015).

expectation in the case of this section, or this song in general. The anticipation of the arrival of the drop is largely caused by the changing texture in conjunction with the meter, as shown in the spectrograms in Example 3.8. The bridge starts with a big dip in the sonic density of the accompaniment, creating a sonic void by cutting high frequencies. A moderately paced filter sweep subtly and gradually starts to fill the void by brightening the texture, creating a small buildup. This buildup starts to intensify significantly at 2:45. The filter sweep becomes much more pronounced, further brightening the texture at an accelerated pace. The kick drum intensifies (along with the snare) by doubling its pace. At around 2:52, additional vocal lines are added in the background. Finally, concurrent with the arrival of the final chord of the loop (ii), the drum texture is significantly intensified by doubling the pace of the kick drum and quadrupling the pace of the snare just prior to the arrival of the drop (3:00).

All of the above play a much more prominent role than the chords in generating anticipation. The main contribution of harmony to the feeling of anticipation in this excerpt is its role as a temporal marker. Each change of chord alerts us to the arrival of a new hypermetric beat. Following a V with a ii would be quite unusual in most tonal music (although not unheard of in rock music), but in “The Night is Still Young,” it is part of the style. The V is only a V, not a dominant. The moment in which the feeling of anticipation is greatest occurs when we hear the ii. Does this mean that ii is the “dominant?” Hearing ii in this way would be consistent with the idea discussed earlier that 20th-century rock and pop music have redefined the dominant, and that it no longer has to be V. However, it would be difficult to hear anything in the pitch content of this ii that would lead to a conclusion that it has anticipatory properties similar to that of a traditional dominant. The “dominant function,” in this case, is carried out primarily by the filter sweep, the drum intensification, and the added vocal texture, combined with the meter.



Example 3.8. Nicki Minaj, “The Night is Still Young” (2015), 2:30-3:05.

Tonal Ambiguity

In common-practice tonal pieces composed in minor keys, composers typically establish the tonal center by introducing the leading tone very early in the piece, thereby creating a clear distinction between the minor key and its relative major. The VII and III triads are normally used at moments when there is deliberate tonicization of the relative major, or as pivot chords in modulations. In contemporary pop music, on the other hand, songwriters generally avoid tonal tendencies, including raising $\hat{7}$ in minor keys to generate a leading tone. As a result, the distinction between the tonal center of a minor key and its relative major can become fluid and ambiguous.

An example of this can be heard in Ariana Grande's "Problem." Based on a melody, that revolves around a G \sharp -minor triad, as well as a G \sharp drone over which the choruses and Iggy Azalea's rap verse are sounded, it is safe to conclude that this song is in the key of G \sharp -minor. However, this conclusion is based purely on the substantial presence of the G \sharp -minor triad in the song, and not on a cadential confirmation of G \sharp as the tonal center. It is completely plausible to hear each individual moment in the song in B-major and to interpret the G \sharp -minor triad as vi. There are no leading tones or any other tendency tones that require us to hear it in G \sharp -minor. In fact, the only instant that resembles a traditional cadential progression (due to the chord progression, not the melody) points to B-major as the tonic triad.

Let us examine this: Example 3.9 shows a melodic transcription of the first verse (0:11-0:30), along with the letter representation of the accompanying triads. The 16-measure melody is structured in **aaba** form. The **a**-sections are accompanied by an E-F \sharp -G \sharp m loop, which would be interpreted as IV-V-vi in B-major in traditional tonal music, but can be idiomatically heard as VI-VII-i in G \sharp -minor in most Western popular music. However the **b**-section replaces the G \sharp m chord with a B major chord, altering the loop to sound as a IV-V-I progression in B-major. The melody of **b** also emphasizes the tones of a B-major triad. But does this necessarily mean that the entire verse can only be interpreted in G \sharp -minor with a momentary tonicization of B-major? To my ears, it is equally plausible to hear the E-F \sharp -G \sharp m loop as IV-V-vi in B-major or as VI-VII-I in G \sharp -minor.

a E F# G#m
Hey ba-by e-ven though I hate you I wan-na love you I want you _____

a E F# G#m
And e-ven though I can't for - give you I real-ly want to I want you _____

b E F# B
Tell me tell me ba - by _____ Why did you leave me _____

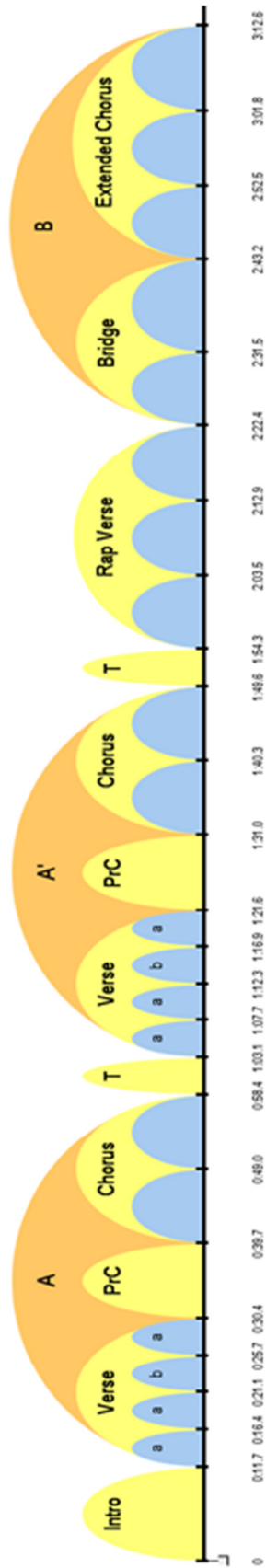
a E F# G#m
'Cause e-ven though I should-n't want it I got-ta have it I want you _____

Example 3.9. Ariana Grande, “Problem” featuring Iggy Azalea (2014). Transcription of first verse (0:11-0:30).

The ambiguity continues throughout the song. The pre-chorus (0:30-0:40) follows with the same chord loops - E-F#-B and E-F#-G#. This is repeated in the bridge as well (2:22-2:43). The harmonic backdrop to the rest of the sections in the song (intro, choruses, and Iggy Azalea’s rap verse) is simply a G#-minor drone (see Example 3.10 for form breakdown). None of these chords contain a leading tone that would strengthen the feeling of G# as the tonal center. Moreover, each melodic phrase in this song contains the note F#, the natural $\hat{7}$ in G#-minor and, more importantly, the dominant scale degree in the key of B-major.

Similarly to “The Night is Still Young,” the lack of commitment to the tonal center as a goal toward which the music gravitates allows other sonic aspects, such as changes in texture, filter sweeps, and drum patterns, to take the lead in creating anticipation. While the chord loops do participate in delineating the form by changing between sections (the verse and pre-chorus have similar chord loops, but they are not identical), changes in texture, along with the vocal melody and lyrics, play a more significant role. In general, Grande’s songs may tend to slightly deviate from the new norms, by using some tonal tendencies and varying the chord loops between sections, but they still largely avoid cadential moves and rely on sound production techniques for important structural anticipatory moments.

As I mentioned earlier in this chapter, pop music is an eclectic genre that routinely borrows from other styles. A big advantage of retaining the diatonic collections and root-position triads as the basis for the pitch content is the potential for activating traditional tonal mechanisms, which is occasionally realized, to varying degrees. Nevertheless, even when pop songs incorporate aspects of traditional tonal models, as is the case in “Problem,” they still allow timbre, gesture, and spatialization to play the lead role in the musical narrative.



Example 3.10. Ariana Grande, "Problem" (2014).

CONCLUSION

This chapter begins to explore the symbiotic relationship between sound production and tonality in contemporary pop music. Traditional tonal syntax and the newly-emergent sonic syntax, which originates in electronic dance music, rely on entirely different apparatuses, but share a common underlying idea – tension and release. While today’s pop music retains tonality and many of its fundamental objects (e.g., tonal centers and triads), the role of tonal harmony has evolved to minimize its syntactical dominance, by nearly eliminating the use of mechanisms that increase the feeling of tension and anticipation, such as cadential progressions, “activated” tendency tones and chromatic alterations. This evolution paves the way for sound production to assume a lead role in guiding the listener through a song’s musical narrative using a range of techniques that include timbral gestures, spatial manipulations, and alterations of sonic density and energy. These techniques will be explored in depth in Chapter 4.

CHAPTER 4

THE ANTICIPATION CHAIN: CREATING TENSION AND RELEASE BY MANIPULATING SONIC DENSITY, SONIC ENERGY, AND RHYTHMIC INTENSITY

As I discussed in Chapter 3, this new sonic syntax features some similarities to tonal syntax. Tonal harmony relies on the tonic-predominant-dominant-tonic cycle. The sonic apparatus in contemporary pop music (as well as in EDM, where it originated) operates on a partially analogous cycle: *setup-buildup-peak*. All three of these sonic functions are anticipatory, to varying degrees. The setup, as its name would suggest, sets up an expectation for a buildup. The buildup builds up tension toward the peak/climax. The peak represents a high level of tension waiting to be released. Unlike the harmonic cycle, however, the sonic cycle does not contain a closing function. The peak section is typically followed by a new setup or buildup, continuing the process of intensifying the music toward an additional peak.

A song starts with a baseline energy level, usually in the first verse, which serves as a kind of never-again-achieved equilibrium. The tension is then increased toward a climax, either by increasing the sonic energy *or* by abruptly decreasing it (if we fall below a level of energy

previously experienced, we anticipate a return to the higher level). When a climax is reached, the sustained high level of energy tenses us up in anticipation of release. The formal trajectory of a song is dictated in part by such manipulations of sonic energy at various points in a song.

This chapter focuses primarily on techniques associated with the buildup function. In general, the setup and peak are sections or subsections within the sonic cycle that represent relative levels of sonic energy – the setup, normally at the beginning of a cycle, is relatively low in sonic energy, and the peak represents a climax of sonic energy. The two are fairly easy to distinguish by these parameters, in most contexts. The buildup function, however, is created by a range of techniques that at times can appear to be contradictory, such as the aforementioned situation in which incrementally adding sonic layers or abruptly removing them both increase tension (to be more explicit – the buildup function refers to a buildup of tension, and not necessarily a buildup of sonic energy). Functionally, the techniques detailed in this chapter can be part of a prolonged section that functions as a setup, but at a self-contained level they all function as buildups. These structural relationships are discussed in detail in Chapter 5.

TECHNIQUES

There are numerous techniques for creating tension in pop music, but they all rely on three gestural principles. The first is *ascending sonic motion*, which directs the listener toward a climax. The second is *creating a sonic void*, which causes the listener to yearn for sonic activity in a certain part of the spectrum. The third is *intensifying rhythmic activity*, which is the rhythmic equivalent of ascending sonic motion. Ascending sonic motion can be achieved by adding sonic layers in distinctive steps, or by using a linear gesture, such as a filter sweep or a pitch bend. A

sonic void is created by vacating an area in the sonic spectrum (or muting the entire spectrum), usually abruptly. In the following pages, I will provide examples of the most common anticipatory techniques, including gradual ascending gestures (i.e., filter sweeps and pitch bends), step-by-step alterations of sonic energy (including filling and vacating sonic spaces), and intensifying rhythmic activity.

Gradual Ascending Gestures (filter sweeps and pitch bends)

Gradual ascending gestures are perceived as goal oriented due to their directional character. Because that direction is upwards, the anticipated goal is expected to constitute an energetic peak. The two most common gradual ascending gestures are the filter sweep and the pitch bend. The following will detail their characteristics and provide examples of their use in contemporary pop songs.

Filter Sweep Risers

Filter sweeps are by far the most common way by which gradual ascending motion is expressed. In particular, I am referring to ascending filter sweeps in which an equalizer is used to cut high frequencies and then gradually release them over a period of time. This produces a sweeping gesture that is perceived as ascending linearly. The combination of the gradual motion and the unchanging meter and hypermeter induces a very strong feeling of anticipation. The fixed hypermeter makes the timing of arrival extremely predictable and the directional nature of the sweep points toward an imagined goal, which can be anticipated by the listener. The filter sweep can be applied to a single sound (often to noise), to a group of sounds in the mix, or to the entire mix. It can function as the central gesture in a big, climactic buildup, or as a small gesture at the end of a section that helps to emphasize the arrival of the next.

A moderately-paced filter sweep progressing over the duration of an entire section or subsection is often the centerpiece of a buildup toward a climactic drop, often referred to as a *riser* in pop jargon. One such riser, which I discussed briefly in chapter 2, occurs at the second half of the bridge in Britney Spears's "Till the World Ends" (2011, 3:07-3:22). In this excerpt, the filter sweep is applied to the entire mix, including the vocals, drums, bass, synth sounds, and noise tracks. Every element in the mix is "climbing" up in the same direction, gradually accumulating high frequency content, which produces a very focused and impactful goal-oriented gesture. The filter sweep is then interrupted by a break on the last beat prior to the drop, heightening the level of anticipation by creating a sonic void. In this particular excerpt, the filter sweep is not only the central gesture in the buildup, it *is* the buildup, since it is applied to every single sound in the mix.

An interesting feature of the relationship between the buildup and the drop in this riser is that, prior to the break, the low frequency content had not been removed or obscured as preparation for the drop. As I mentioned previously, buildups tend to temporarily remove the low frequency content in order to create a yearning for its return in the form of the "dropping" of the drum and bass. In "Till the World Ends," the drum and bass achieve near-maximum presence early in the buildup, since the ascending filter sweep releases low frequencies first, and this presence is maintained until the break. Thus, instead of relying on a sonic void to be created during the buildup, the sense of anticipation is generated primarily by: a) the general accumulation of sonic density, which reaches its climax at the drop with an explosion of noise that saturates the sonic spectrum; b) directing the listener's attention to the higher frequencies as they are progressively amplified, without reducing the presence of the low frequency content; and c) the sonic void (all of the sounds removed except for the vocals) created at the break.

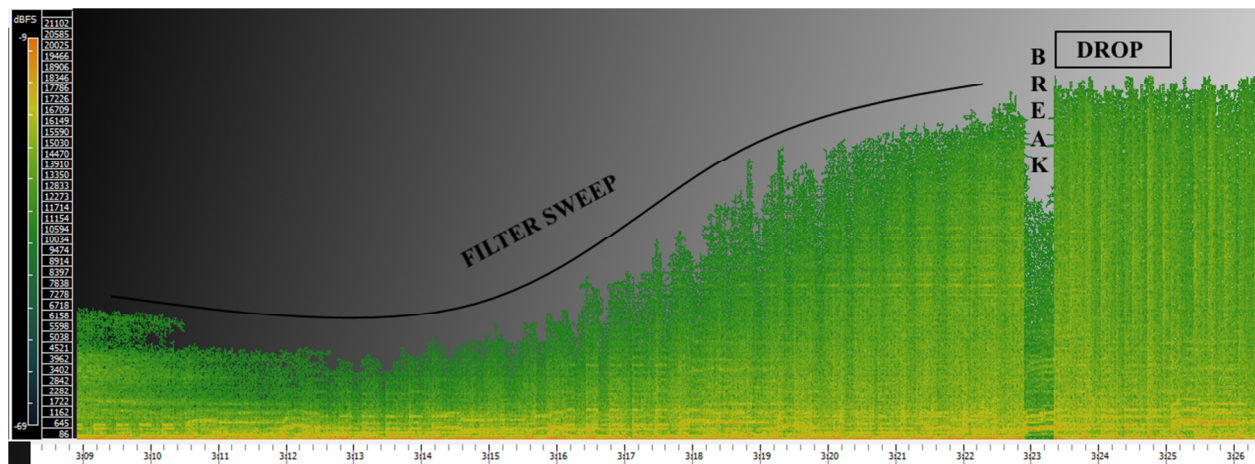
By contrast, Alesso's "Heroes" (2014, featuring Tove Lo) contains a buildup (0:53-1:08) that produces a similar effect, but uses different means. Here, too, a filter sweep is the central component of the buildup, but instead of being applied to the entire mix, it is applied to a semi-pitched noise track that directs the listener's attention in an ascending motion. Concurrently, another filter sweep is applied to a repetitive synth riff that becomes increasingly pronounced as the buildup progresses due to the gradually augmented presence of high frequency content. As the imminent drop approaches, the low frequency content of the drums and bass is steadily obscured until it almost disappears. This creates a void which is filled upon the drop's arrival (1:08) - a forceful return of the kick drum and bass, supporting a climactic section in which the synth riff becomes the focus of the listener's attention, or the *hook*.

We can see the similarities and differences between these two buildups in examples 4.1¹⁴ and 4.2. In "Till the World Ends," the sonic spectrum is initially almost entirely empty (with the exception of an echo from the "explosion" that triggered the riser), and is gradually filled in a sweeping motion. In "Heroes," the spectrum is relatively full throughout the buildup but we see the noise sweep gradually amplify the higher frequencies (expressed in the spectrogram by green areas transforming into bright yellow, primarily between 1:01 and 1:08). We can also clearly see the low frequency void created prior to the drop in the melodic range spectrogram. In both of those buildups, the listener is gradually led to focus on the high frequency content by the ascending filter sweeps, but the makeup of the two sweeps is different. In one it is a sweep of the entire mix which is followed by an additional, unfiltered iteration of the same music. In the other, it is a noise track that points upwards and a riff that is gradually brought to the foreground, preparing it to be the central hook in the following section. One creates a sonic void by abruptly

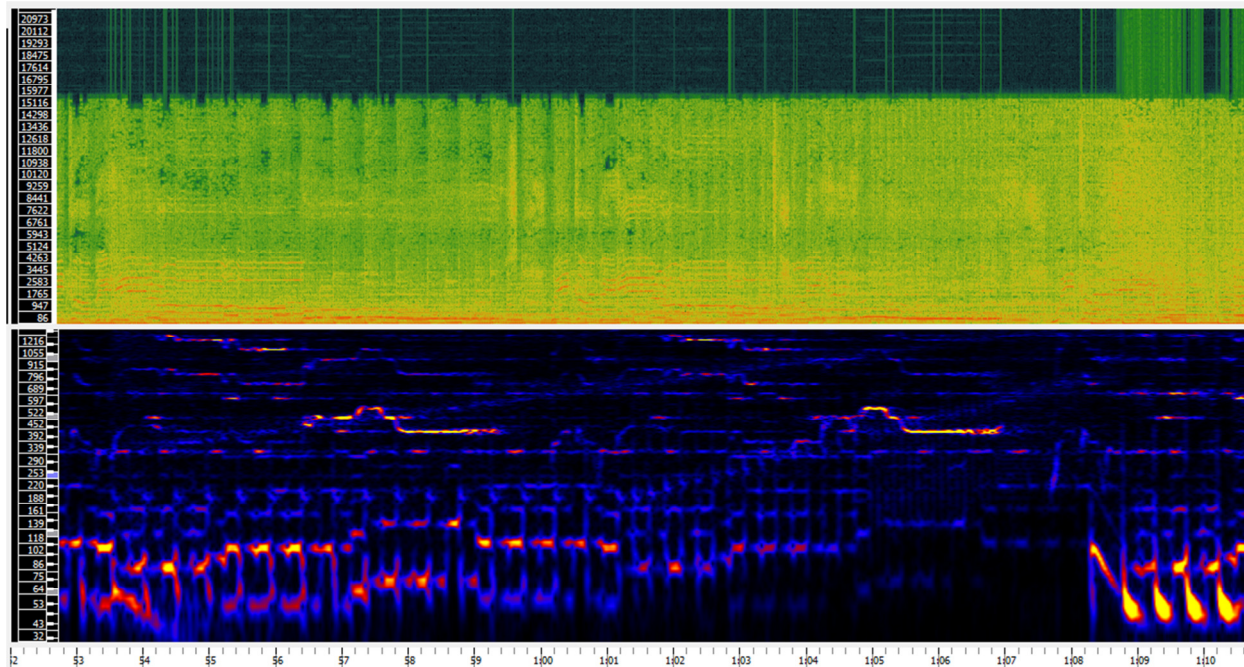
¹⁴ Note: In example 4.1 I used image editing software to remove the visual representation of very low amplitude frequencies in order to better show the gradual amplification of higher frequencies over time.

silencing everything but the vocals, the other by gradually filtering out low frequency content.

The underlying idea is identical in both – a gesture that causes the listener to crave the return of a full texture (particularly the low frequency content) by taking parts of it away. Musical context dictates the differences between these risers.



Example 4.1. Britney Spears, “Till the World Ends” (2011), 3:07-3:23. Edited spectrogram.

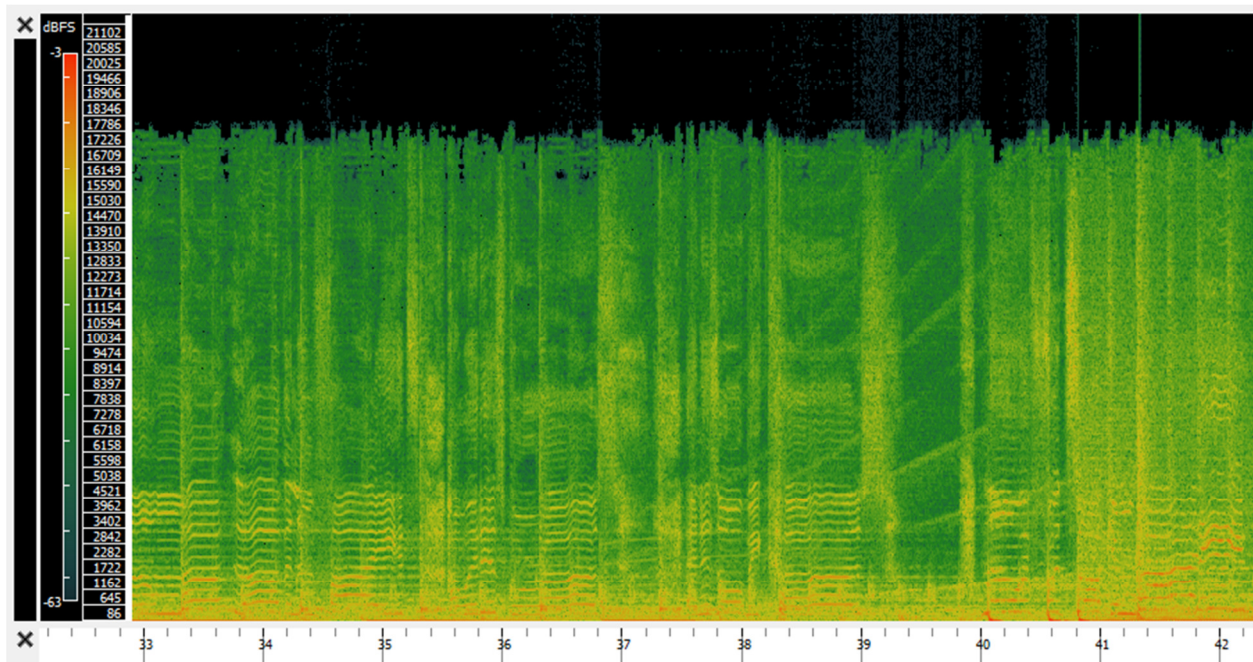


Example 4.2. Alesso, “*Heroes*” featuring Tove Lo (2014), 0:53-1:08.

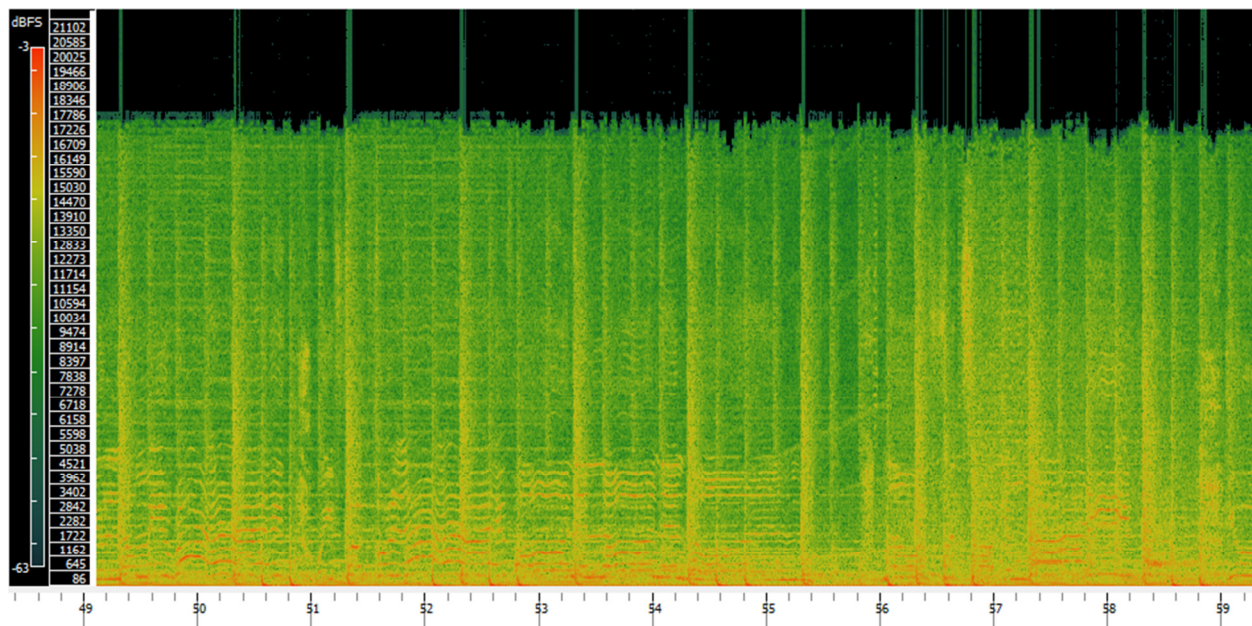
Small Ascending Sweeps

The above examples feature filter sweeps that build up over the course of an entire section, but filter sweeps also have anticipatory qualities when they come in the form of small gestures that signal the arrival of the next section. In Tove Lo’s “Talking Body” (2014), for example, a partially-pitched noise sweep, similar to the one in “Heroes,” starts to ascend at 0:37, gesturing toward the arrival of the chorus (0:41). The effect here is similar to that of the big buildups in “Heroes” and “Till the World Ends,” but on a significantly smaller scale. The sonic texture of the pre-chorus is rather light, as indicated by the prevalence of darker green in the mid- and high-frequency areas prior to the chorus (0:41) in the spectrogram (Example 4.3). In this texture, we can clearly see the sweep “slice” through the texture, and indeed it is heard as a central element, drawing our attention to the arrival of the chorus. Midway through the chorus,

which consists of two nearly-identical iterations of the same music, the same sweep is heard (0:52-0:57) as the music gears up for repeating the first part. This time, this sweep is not as distinct as its previous iteration, since the surrounding texture is much denser. It only becomes distinctly noticeable at ~0:55, and is somewhat obscured by the kick drum at ~0:56, just before the downbeat at ~0:57. This aural experience is accurately represented in the spectrograms - While the ascending lines that represent the sweep at the end of the pre-chorus (0:37-0:41) in Example 4.3 are quite clear, they are much harder to detect (0:54-0:57) in the full textured chorus, as shown in Example 4.4.

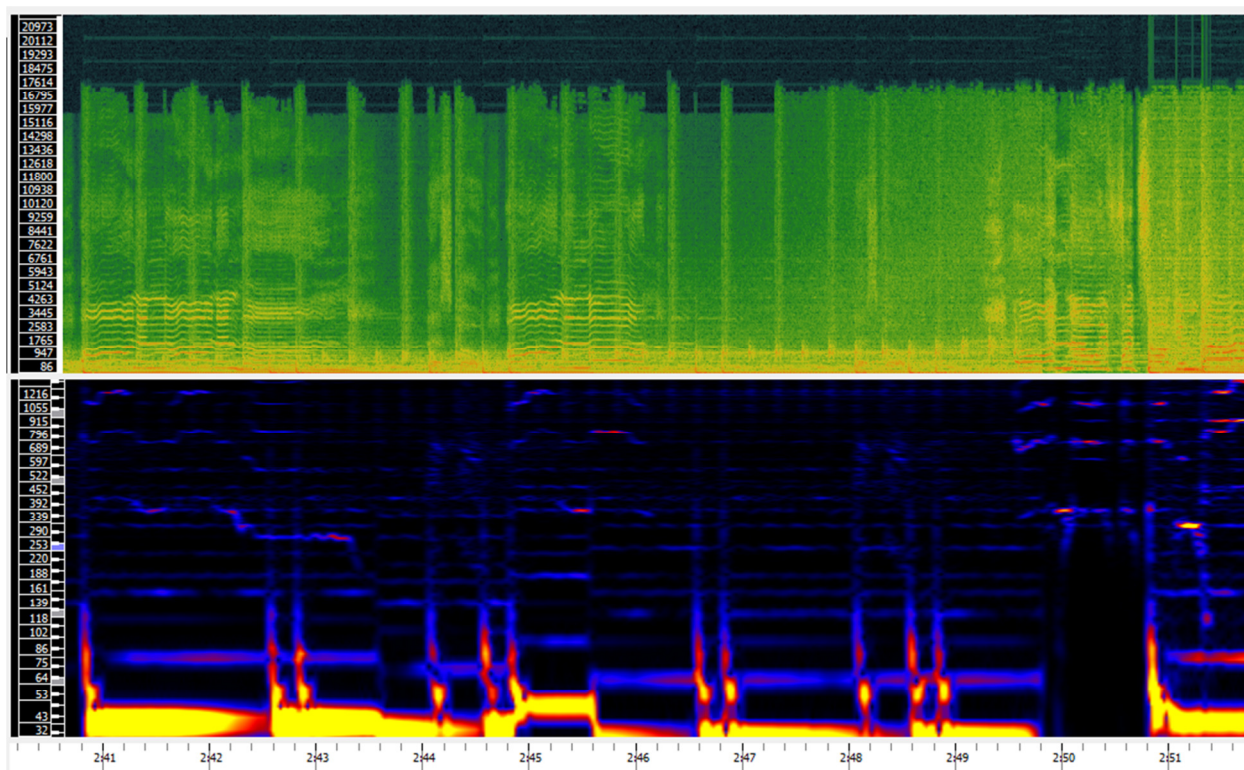


Example 4.3. Tove Lo, “Talking Body” (2014), 0:33-0:42.



Example 4.4. Tove Lo, “Talking Body,” 0:49-0:59.

In the same song, a similarly short filter sweep (2:46-2:51) is heard at the end of the bridge, leading up to the final chorus. This time, there are no distinct lines in the spectrogram that represent the ascending motion (Example 4.5). Instead, we see a gradual accumulation of higher frequencies in more-or-less equal amplitude between 2:46 and 2:50 in the plain spectrogram. This is because the filter sweep in this case is applied to a white noise track, rather than to partially-pitched noise. This filter sweep is somewhat more pronounced than the previous two, in part because of the break at 2:50, which heightens the tension, and in part because it occurs at a structurally more climactic moment. Nevertheless, it functions in the same way, playing a significant role in heightening the listener’s feeling of anticipation toward a predictable outcome – in this case, the final chorus.



Example 4.5. Tove Lo, “Talking Body,” 2:41-2:52.

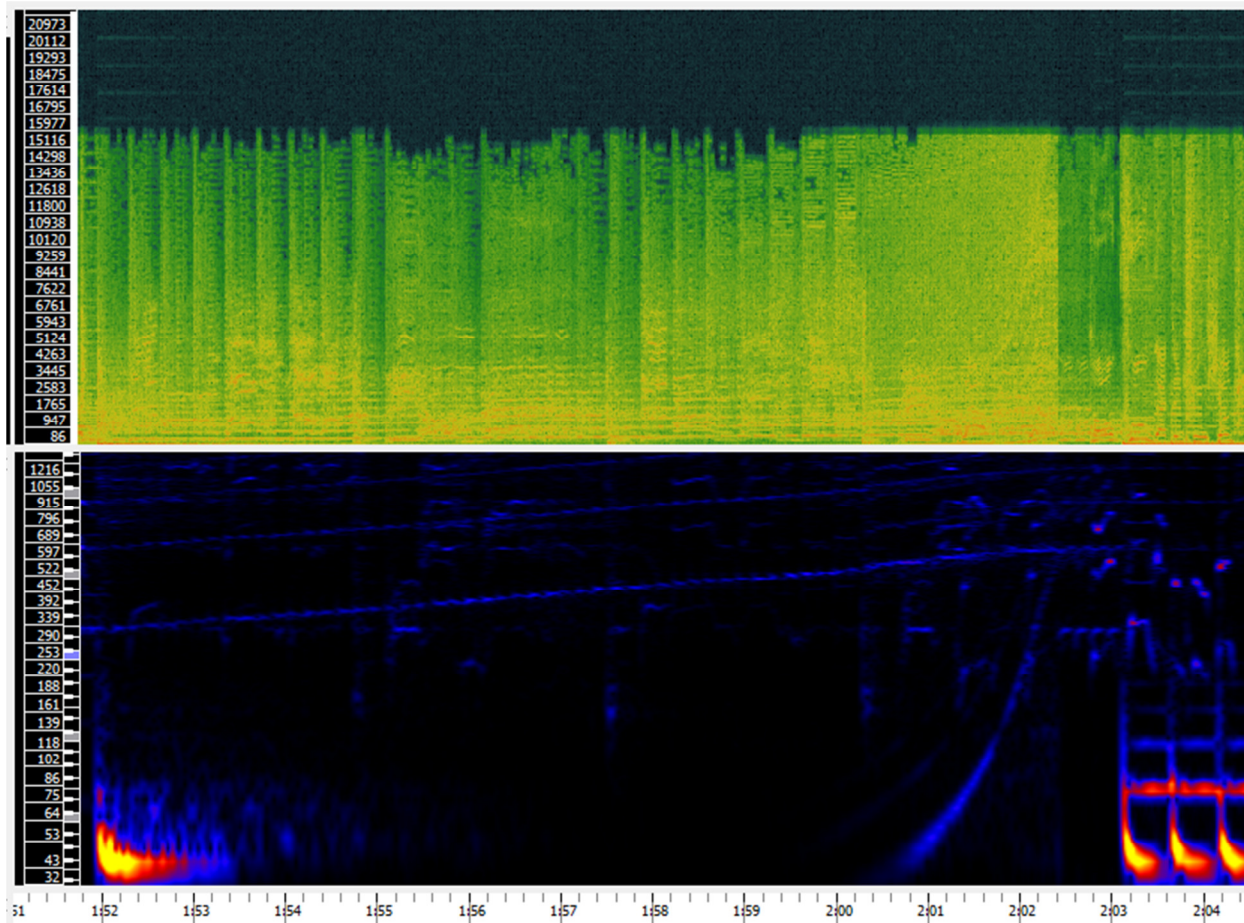
These short filter sweeps occur in numerous songs. While big filter sweep risers are usually reserved for heavily dance-oriented songs, short filter sweeps signaling the arrival of a new section are extremely common, even in pop ballads. A few of the songs that contain such gestures, some as short as one second, include Ariana Grande’s “Break Free” (2014), at 0:35-0:37, 0:50-0:52, and 1:42-1:44; Britney Spears’s “Hold it Against Me” (2011), at the change of every major section (for example, 0:06-0:07, 1:03-1:05, and 1:39-1:41); Taylor Swift’s “Wildest Dreams” (2015), at 2:33-2:38 and 2:52-2:53; and Tove Lo’s “Not on Drugs” (2014), at 2:10-2:13, 2:17-2:21, and 2:55-2:59 (which leads to an abrupt end of the song). These short filter sweeps are the “little siblings” of the big risers – they too provide a goal oriented gesture that

leads the listener to anticipate an arrival of a new section on a hypermetric downbeat, but they do so on a more local structural level.

Ascending Pitch Bends

Ascending pitch bends (*glissandi*) are very similar to filter sweeps in their effect. They also point upwards and draw our attention to the upper parts of the sound box, and are often followed by a contrasting drop. Although their focused presence (due to the pitch occupying very focused frequency “lines”) often positions them in the foreground of a buildup, pitch-bends are not commonly used as stand-alone anticipatory gestures, and are frequently paired with filter sweeps to enhance a riser. One example of this has already been presented in chapter 3 – the riser in Rihanna’s “We Found Love” (2011), in which a rhythmically broken pitch bend, while clearly in the foreground, is one of several anticipatory gestures (along with noise sweeps, drum intensification, and an additional unbroken pitch bend in the lower part of the spectrum) that lead up to a climactic drop.

Another previously mentioned song, David Guetta’s “Hey Mama,” features a buildup (1:52-2:03) in which a pitch bend does initially stand on its own, but is joined by a noise sweep (1:58-2:02), followed by a break (2:02-2:03), just prior to the arrival of Nicki Minaj’s rap verse, which begins with distinct bass and kick drum attacks. Example 4.6 shows the ascending lines that represent this pitch bend, gradually leading the listener toward the climax (in both spectrograms).



Example 4.6. David Guetta, “Hey Mama” featuring Nicki Minaj (2014), 1:52-2:04.

Unlike the filter sweep, which occurs regularly in a number of contexts, it is uncommon for a pitch bend to serve as a small signaling gesture between sections, and it is reserved primarily for risers. I can only speculate as to why this is the case. Perhaps producers feel that a short pitch bend at the end of a section would compete with the vocals for the listener’s attention, while filter sweeps and breaks better complement the vocals in this context. Nevertheless, the pitch bend’s function, like the filter sweep in its various forms, is to guide the listener toward a goal in a sweeping, ascending motion.

Adding and Removing Sonic Layers

Intuitively, we tend to assume that if one action produces a certain result, the opposite action will produce the opposite result. This is not the case when the texture of a song is altered by adding or removing sonic layers. In most contexts, both of these actions increase tension, though for different reasons. Adding sonic layers increases the sonic energy, and therefore leads an ascent that we anticipate will eventually reach a climax. Removing layers, on the other hand, creates a sonic void. While this is a move in the opposite direction, the void causes us to yearn for the lost sounds, thereby increasing our feeling of tension and anticipation for their return. Within the idiomatic boundaries of the typical pop song structure, there is only one instance in which abruptly dropping sonic layers is perceived to decrease tension rather than increase it. This occurs in the “sonic fall” from the first chorus to the second verse. In this moment, it is usually the case that the relief of defusing the climactic energy in the chorus outweighs the sonic void created by abruptly reducing the energy. This is discussed in greater detail later in this chapter, as well as in Chapter 5.

In general, there are four directions in which the texture is typically altered between sections (and sometimes during a section): adding low frequency content, adding mid/high frequency content, removing low frequency content, and removing mid/high frequency content.¹⁵ Due to limitations of scope and a desire to focus on the most common occurrences, I am operating under the assumption that adding sonic layers increases rhythmic intensity and removing sonic layers decreases it, but there are, of course, examples to the contrary. In the

¹⁵ I am lumping mid- and high-frequency content in the same category because they are often added or removed together. There are instances of only mid- or high-frequencies being added or removed, but making this distinction is not essential for the analytical purposes of this chapter.

following paragraphs, I will discuss how each of these four types of alteration contributes to increasing tension and anticipation.

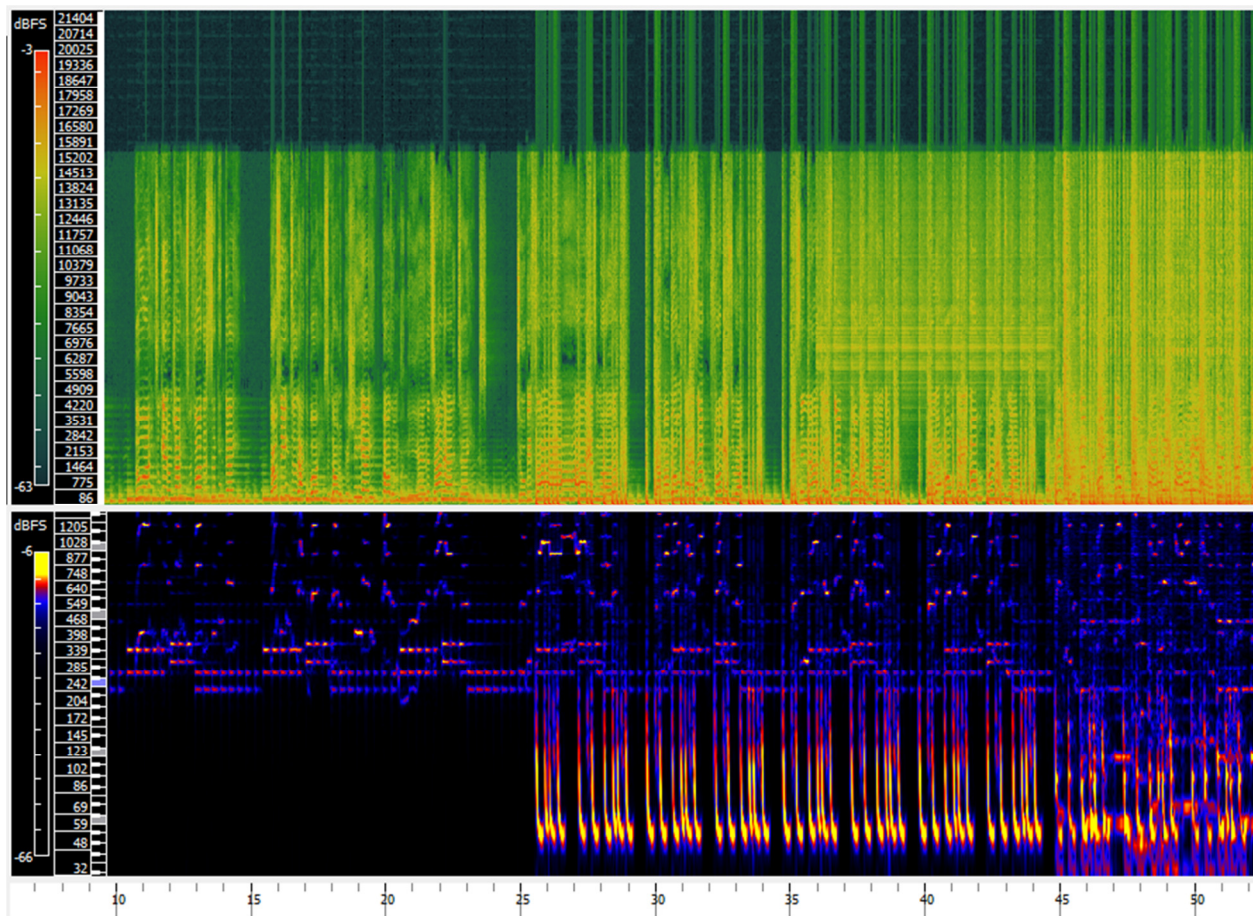
Adding Sonic Layers

In nearly every contemporary pop song, the sonic energy is altered between sections. One example of this can be heard in Daya's "Hide Away" (2015) where the tension is increased by initially adding low frequency content and later adding high frequency content. The texture of the verse (0:05-0:26) consists only of Daya's voice and a simple keyboard accompaniment, both predominantly occupying the middle region of the sonic spectrum, and enhanced with reverb and delay effects in order to "sweeten" the general sound. We can see the vacant low-frequency band leading up to ~0:26 in the melodic range spectrogram (Example 4.7). The kick drum then enters at the beginning of the pre-chorus, adding low frequency presence and signaling to the listener that the sonic energy is moving toward a climax. This is further intensified by background "Hey!" shouts (0:30 and 0:34). Midway through the pre-chorus (~0:36), a repeating open cymbal enters, this time adding presence in the high frequency area (represented by the abundance of yellow above 6000 Hz in the plain spectrogram, 0:36-0:45).

Every step in this development brings the listener closer to the climax in the chorus. It is a well-established norm in the pop genre that once the sonic energy begins to intensify in the pre-chorus,¹⁶ an eventual arrival at a sonic climax in the chorus is expected. Moreover, each step in this type of buildup is expected to occur at the beginning of a chord loop, enabling the hypermeter to guide the listener's expectations. Indeed, we arrive at the peak of sonic energy in the chorus (0:45-1:06). A bass synth, doubled by a sub-bass (represented by the activity below

¹⁶ The section I define as the pre-chorus (0:26-0:45) can be interpreted as a continuation of the verse instead, from a traditional songwriting perspective. However, the sonic developments in this timeframe provide enough contrast for me to interpret it as the pre-chorus. I explore these considerations in more detail in chapter 5.

48 Hz in the melodic range spectrogram starting at 0:45), is added, along with an active hi-hat, and the vocal melody is artificially doubled an octave below. The low-, mid-, and high-frequency regions of the sonic spectrum are all represented in the mix, and there are no sonic voids to be filled. At this point, the “goal” is the release of this accumulated energy, which the listener can anticipate will happen after four cycles of the chord loop, since this was the duration of both the verse and pre-chorus.



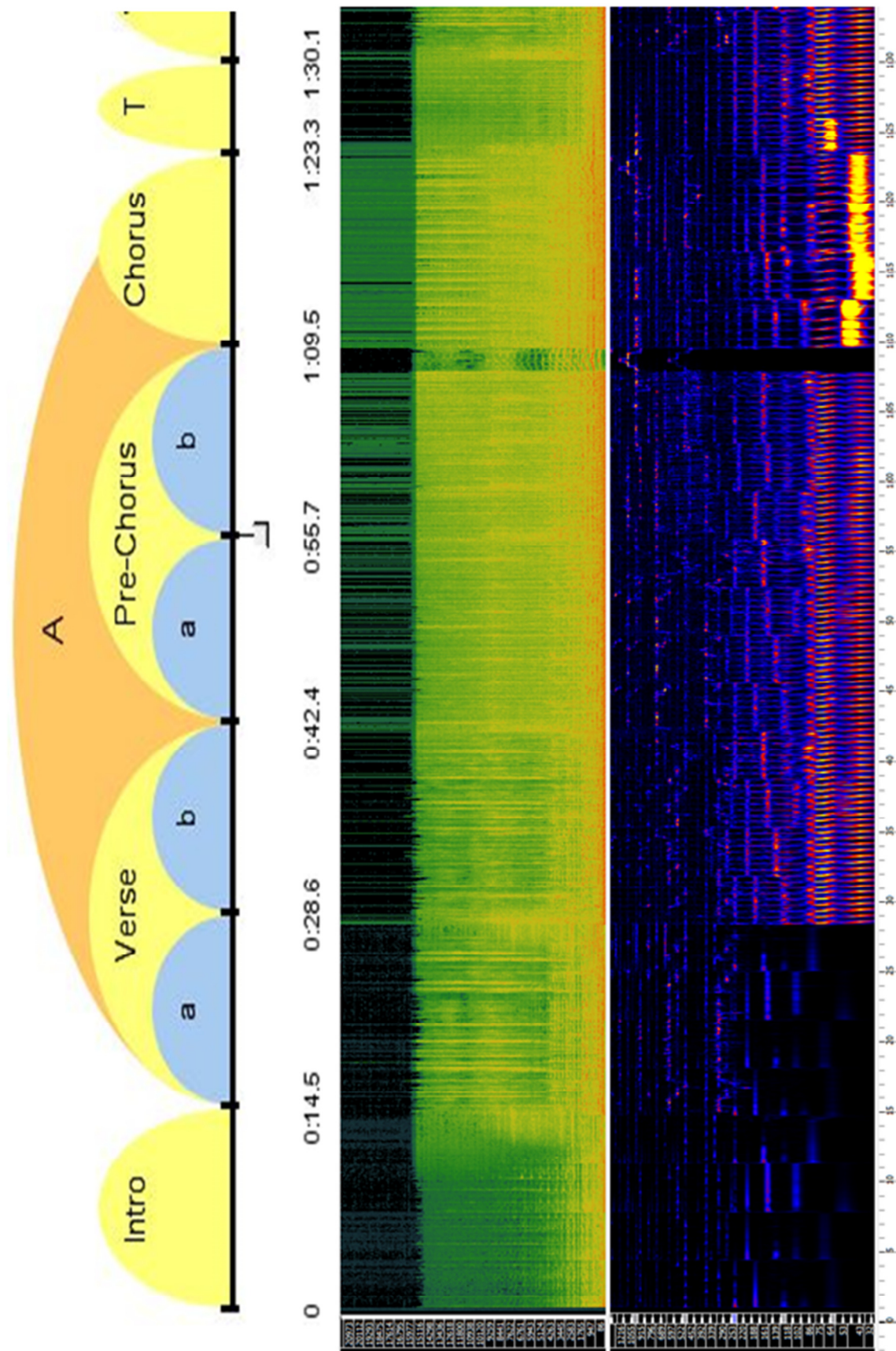
Example 4.7. Daya, “Hide Away” (2015), 0:10-0:52.

“Hide Away” is not an EDM-oriented song, and therefore the climax is not as “explosive” as those we heard following the risers in “Till the World Ends,” “Heroes,” or “We Found Love,” but the fundamental principle is similar. In pop jargon, the arrival of the chorus would not be referred to as a drop, since this term implies a burst of sonic energy following a very intense buildup that features at least one gradual ascending gesture, as we heard in the aforementioned songs. However, although the intensification of energy occurs in distinct steps rather than in a sweeping motion, it is nonetheless a goal-oriented process that aims to arrive at a climax, which makes it in essence a milder version of the EDM-influenced riser. In fact, we can think of each step as a miniature drop, in the sense that it arrives at an anticipated hypermetric downbeat, and fills a sonic void. This is another well-established idiom – the absence of the kick drum (and bass) constitutes a sonic void, even when it has yet to be sounded. The listener anticipates its arrival, not because it was previously present in the song, but because its presence is so prevalent in the genre, that we have a strong expectation that it will eventually arrive.

A similar but more intense process of adding sonic layers occurs in Tove Lo’s “Timebomb” (2015). Example 4.8 shows spectrogram representations of this process aligned with a formal breakdown of it. It is enough to glance at the plain spectrogram to see that the texture of the song becomes progressively denser at fixed, 14-second intervals (or 8 measures, assuming $\text{♩} = \sim 137$). Here, the addition of sonic layers occurs at each new iteration of the recurring chord loop (the intro consists of one chord loop; the verse and the pre-chorus each contain two) rather than the start of a new section. In fact, significant developments occur in the middle of the verse and pre-chorus. The song deceptively starts out very mellow, with a ballad-like piano intro, very light percussion in the mid-range, and a mild noise sweep that resembles a soft cymbal-roll that ushers in the verse. The accompaniment remains identical in verse **a** (0:14-

0:28), but Tove Lo's rhythmically dense vocal line hints that we should expect the accompaniment to catch up.

At the beginning of verse **b** (0:28), the buildup crosses the “point of no return.” The sonic void in the low frequencies is filled by the kick drum and bass, and the rhythmic intensity picks up, generating the expectation that this process will end in a powerful sonic climax. Next, at pre-chorus **a** (0:42) a measured snare-roll enters, adding presence to the mid/high range and further intensifying the rhythm. At the midpoint of the pre-chorus (0:56), a series of high-pitched ascending tones that sound like video game shooting sounds are added. Finally, just before the anticipated arrival (1:07), the sonic carpet is pulled from under our feet and we are left with only Lo's voice. This brings the tension to a peak level. We know that something big is about to happen. And indeed, the “explosion” arrives, in the form of all the primary accompanying sounds lining up with Lo's voice as she sings the line “Bomb, bomb, bomb, bomb” in a powerful, syncopated rhythm.



Example 4.8. Tove Lo, "Timebomb" (2014), 0:00-1:30.

The step-by-step intensification is supplemented by gestures that increase the feeling of anticipation toward each section. Each chord loop unit ends in a progressively more potent anticipatory gesture:

- **Intro to verse a (0:14)** – Noise sweep.
- **Verse a to verse b (0:28)** – Noise sweep + vocal “Aah”s.
- **Verse b to pre-chorus a (0:42)** – Noise sweep + vocal “Aah”s (in this case the same gesture is repeated, and is even somewhat obscured by the full texture, but the presence of the kick drum and bass makes up for this).
- **Pre-chorus a to pre-chorus b (0:56)** – Noise sweep + vocal “Aah”s + tom fill.
- **Pre-chorus b to chorus (1:10)** – Break.

Each addition of a sonic layer increases anticipation toward the next phase, while also participating in the larger process of anticipating the arrival of the chorus/climax. The establishment of the pattern, in which the texture is altered at each iteration of the chord loop, helps to instill confidence in the listener that s/he can predict the next step in the process. The listener does not need to know exactly what is coming, but s/he knows that there is going to be a sonic alteration, and s/he also knows exactly when it will occur.

Although the prolonged process of intensification in this song does not rely on a big linear gesture such as a filter sweep, it nevertheless feels exceptionally climactic. In fact, this is a rather unusual case in which a step-by-step buildup feels as intense as an EDM-oriented riser. This can be attributed, at least in part, to the almost-exaggerated intensity of the succeeding steps. The first two developments (verse **a** and verse **b**) are each more intense than expected. Following the ballad-like intro, Lo’s vocal melody can be perceived as overly dense. Inserting

the bass and kick drum in the middle of the verse may sound surprising and extreme. By pre-chorus **a**, both the sonic density and rhythmic intensity are at extremely high levels, and the addition of the shooting sounds in pre-chorus **b** takes it over the top, causing the listener to feel like an “explosion” is imminent. The break drops the sonic density and rhythmic intensity but significantly heightens the tension. Throughout this buildup, it feels as though we are racing toward something big, and when the chorus is reached, it is simultaneously a satisfying arrival and a further intensification that resolves only at the transition (1:23).

This is perhaps a good time to note that “Timebomb” is an exception to the norm because it includes cadential closure. The chord loop changes halfway through the pre-chorus, leading towards the chorus with a IV. The chorus arrives on what sounds like a cadential 6/4 chord. While this chord does not immediately resolve, it steps back to I⁶, and proceeds to IV, which finally lands on I on the downbeat of the transition to the second verse. Here we have a rare instance in which all of the elements “agree” on closure: the chord progression resolves to I; the vocal melody ends on $\hat{1}$; and the high tension caused by the saturated texture in the chorus is released. This cadence is particularly interesting because it highlights the differences in the expression of the idea of tension and release in common practice tonality versus sound production driven gestures. While the resolution of the buildup occurs at the arrival of the chorus, the tonal resolution takes place after the chorus, at the downbeat of the transition. The cadential closure is only possible because all of the different elements (melody, harmony, sonic texture, and meter) are in sync.

We can see that there are commonalities between the stepwise sonic development exhibited in “Hide Away” and “Timebomb” and linear gestures such as the filter sweep. Both build up intensity toward an expected climax, directing the listener’s attention “upwards,” one to

the high region of the frequency spectrum and the other to the peak of sonic energy. The main difference is that the stepwise process is primarily used for building up tension over a prolonged period of time encompassing several sections, while the filter sweep is more effective for quickly ramping up tension, usually over the course of one section, or as a small gesture leading from one section to another.

Abruptly Removing Sonic Layers

If we were to visualize the ascending sonic gestures, we could imagine a staircase (step-by-step intensification) or a ramp (sweeping ascending motion). Inversely, when sonic layers are abruptly removed, we can imagine a surface being pulled from under our feet, leaving us hanging in the air. The sonic energy drops, but the feeling of tension increases as we are still mentally holding on to the lost energy, desperately hoping for its return before we fall. The other effect of this is the expansion of the dynamic spectrum. All other things being equal, the eventual climax will feel more satisfying compared to an identical climax reached without first falling to a lower level of energy, because it fills a larger sonic void.

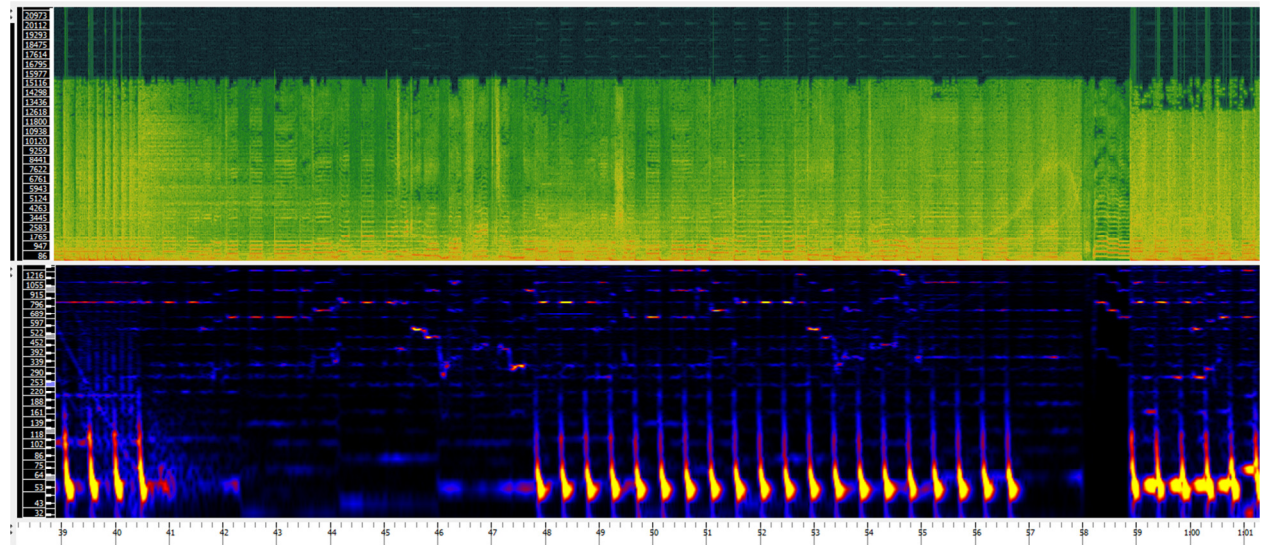
Abrupt removal of sonic layers often occurs at transitional moments or sections, such as the pre-chorus and the bridge, although this is not a hard rule, since contemporary producers increasingly incorporate transitional gestures within sections that are not traditionally considered transitional. The layers that are removed can be either low frequency (e.g., bass and kick drum), or mid/high frequency (e.g., synth pads, percussion). Another common context in which sonic layers are removed is a break, in which all sonic activity comes to a halt (usually two or four beats prior to the hypermetric downbeat), often leaving only the vocals floating in the sonic space.

Both types of abrupt layer removal can be heard in Britney Spears's "I Wanna Go" (2011). This song begins with a step-by-step buildup of sonic energy from the intro (0:00-0:08) to the verse (0:08-0:30) to the pre-chorus (0:30-0:41). Toward the end of the pre-chorus, a drum fill and a vocal "aah" that gradually becomes more audible lead us to believe that we are about to arrive at a climactic section. But instead, as we can see in Example 4.9, the drums and lead synth drop out, replaced by a soft pad accompanying Spears's voice. This unexpected drop in energy puts us in a state of anticipation, waiting for the music to re-intensify. Moreover, between 0:41 and 0:48, the accompanying synth pad is side-chained.¹⁷ In this case, the synth pad is triggered by a completely muted kick drum, which means that we hear the sonic space being created, though the kick drum is not heard. The void created by the absence of the kick drum is, as a result, extremely emphasized.

The kick drum returns to fill the void at 0:48, but the climax is still only on the horizon. However, the returning kick drum triggers a mini-riser, headed by a moderately-paced noise sweep. This mini-riser does not reach its climax at the end of the 4-measure hypermetric unit (for the purpose of this description, I am assuming that the kick drum attacks represent quarter notes in a 4/4 meter). Instead, the tension is heightened by a number of things: the mini-riser being extended by two measures (0:55-0:59), the noise sweep beginning to accelerate, and the break. The break comes in two parts. First, the kick drum drops out again at the beginning of the second measure of the extension. Next, all other sounds come to a halt (0:58), completely vacating the sonic space for Spears to sing the pickup to the repetition of the chorus (0:59-1:17), this time a "real" climactic one. This example illustrates the unique quality of the break – it brings with it

¹⁷ Side-chaining is a process in which an attack of sound x, usually the kick drum, triggers sound y, usually an accompanying synth or a group of accompanying sounds, to be momentarily compressed to near silence. This creates more room for sound x in the sonic space, and it also makes for a more 'bouncy' accompaniment, which is usually desirable in this genre.

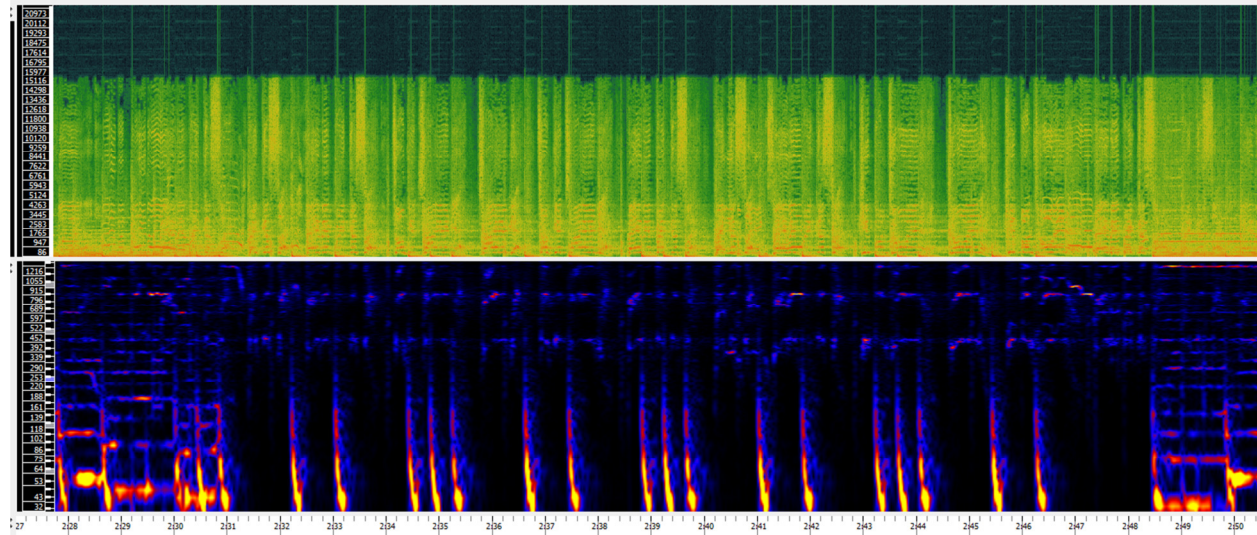
perhaps the greatest degree of inevitability felt by the listener that the climax is about to arrive. We anticipated an arrival at the end of the hypermetric unit in 0:48-0:55, and our feeling of tension increased when what we predicted did not materialize, but the hypermeter alone does not generate as strong a feeling of inevitability as the break does.



Example 4.9. Britney Spears, “I Wanna Go” (2011), 0:35-1:01.

While the primary factor in creating tension in “I Wanna Go” is the removal of the drums, it is the other way around in Becky G’s “Break a Sweat” (2015). Instead of cutting out the drums, a similar effect is achieved by cutting everything *but* the drums and the vocals (2:31-2:49). Here, there is no filter sweep to build back up to the climax. The abrupt emptying of much of the sonic spectrum (see Example 4.10) is sufficient to cause us to anticipate a return to a full texture. As a result of the kick drum pattern combined with the otherwise empty low-frequency region, there is a naturally occurring break at the end of this section, but its effect is not as

pronounced as that of the break in “I Wanna Go,” because it comes in the context of an already-vacant texture.

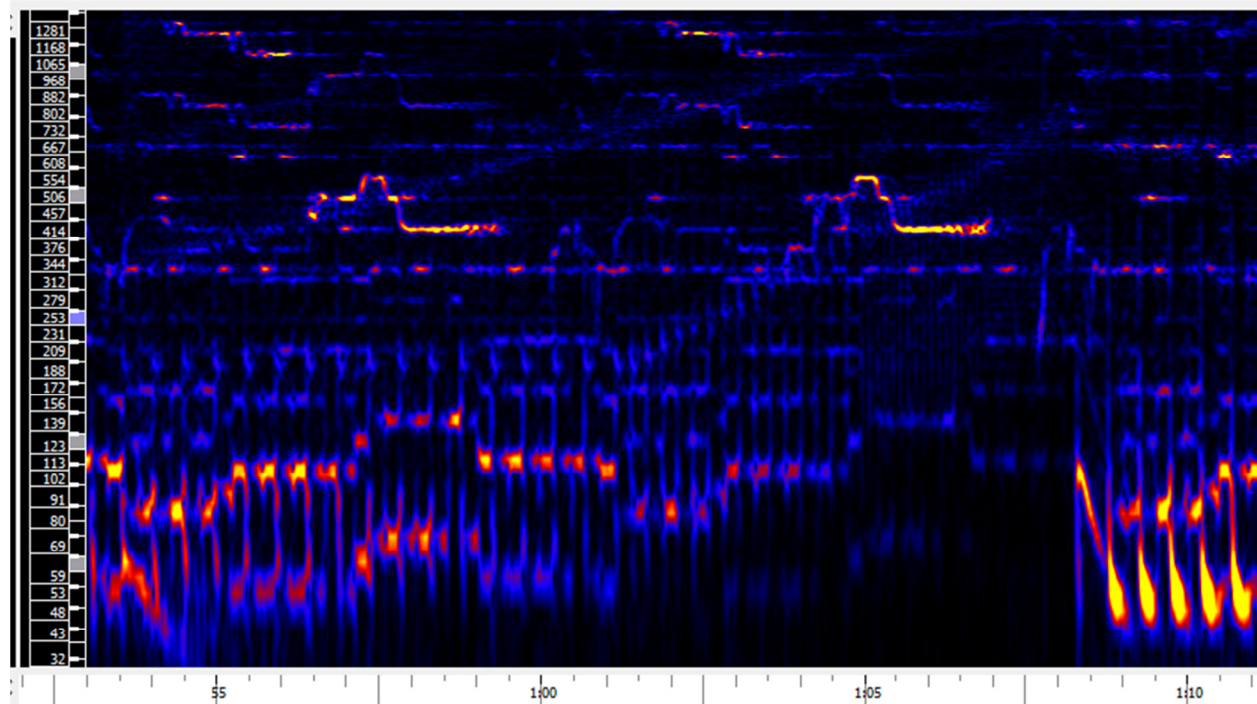


Example 4.10. *Becky G, “Break a Sweat” (2015). Only vocals and drums are present between 2:30 and 2:49.*

Intensifying Rhythmic Activity

One staple of EDM is the accelerating drum part at the end of the buildup. Returning to the buildup in Alesso’s “Heroes” (0:53-1:08), the kick drum and snare at first play quarter notes in rhythmic unison (the kick drum is initially more prominent in the mix, but the snare gradually takes over via EQ manipulation). Halfway through the buildup (1:01), the drum hits accelerate to eighth notes. Two measures later, they turn into sixteenth notes, and then into sixteenth-note triplets in the final measure of the process. As this process unfolds, the kick drum is gradually filtered out and the snare drum is progressively transposed upwards, to go along with the

concurrent filter sweep (see Example 4.11). This is an idiomatic gesture, which occurs in numerous songs (in varying forms), including Ariana Grande’s “Break Free” (0:51-1:14), Pittbull and Ne-Yo’s “Time of Our Lives” (2014, 2:51-3:07), Will.I.Am’s “Scream and Shout” featuring Britney Spears (2013, 3:27-3:42), LMFAO’s “Party Rock Anthem” featuring Lauren Bennett (2011, 3:01-3:17), and Zedd’s “Addicted to a Memory” (2015, 1:21-1:36). We can also hear miniature versions of this gesture in “Break Free” (0:37-0:51), Katy Perry’s “Teenage Dream” (2010, 1:40-1:57), and Becky G’s “Lovin’ So Hard” (2015, 0:33-0:46).



Example 4.11. Alesso, “Heroes” featuring Tove Lo. The rhythmic activity intensifies and the snare drum is gradually transposed up, while low-range sounds (kick drum and bass) are filtered out in anticipation of the drop (1:08).

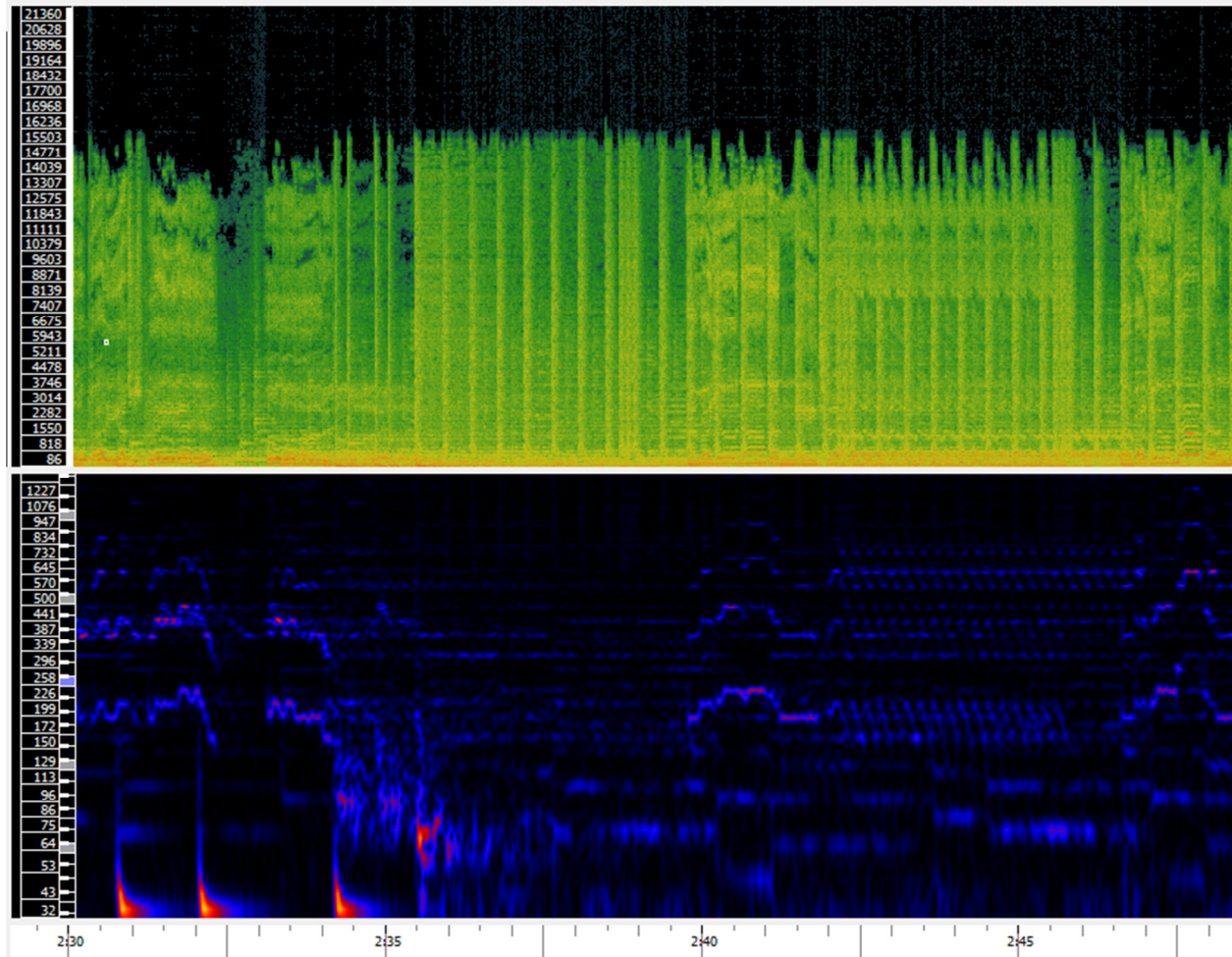
The intensification described above is the rhythmic parallel of the filter sweep – its most intense form is typically used in risers, and its miniature form signals the end of a section and the arrival of a new one. Accordingly, there is also stepwise addition of rhythmic layers, which mirrors the stepwise addition of sonic layers and incrementally increases tension. We can hear this in Taylor Swift’s “We are Never Ever Getting Back Together” (2012, 0:00-0:40). This song begins with an intro that consists solely of an acoustic guitar riff. The verse (0:06-0:28), in which Swift begins to sing, retains this riff, and adds a steady kick drum beat in quarter notes, which starts to propel the energy forward. Mid-verse (0:17), a hi-hat starts to double the kick drum, which adds a subtle sonic layer, but the steady quarter notes persist. The kick drum continues in quarter notes throughout the pre-chorus (0:28-0:40), while the hi-hat accelerates to sixteenth notes, ramping up the sonic energy without adding a new sonic layer.

Every step so far has increased the tension toward an imminent climax by using rhythmic intensity in a manner similar to the step-by-step addition of sonic layers that we saw in previous examples. Instead of a timbral sequence of events, we hear a (mostly) rhythmic sequence: No drum beat → quarter notes kick drum → added hi-hat → sixteenth-notes hi-hat. Only at the very end of this process (0:37-40), a short noise sweep is added to “seal the deal” and heighten anticipation toward the chorus.

Jack Ü’s “Where are Ü Now” featuring Justin Bieber (2015) includes a rhythmic buildup that is a hybrid of the stepwise buildup shown above and rhythmic intensification common in buildup-drop sequences. Following a drum-less second verse (2:07-2:34), a steady sound resembling a fusion between a handclap and a snare drum (to which I will henceforth refer as a snare-clap) is introduced in the pre-chorus (2:34-2:58). The snare-clap (represented by yellow vertical lines in the plain spectrogram in Examples 4.12.a and 4.12.b) carries a mostly-steady

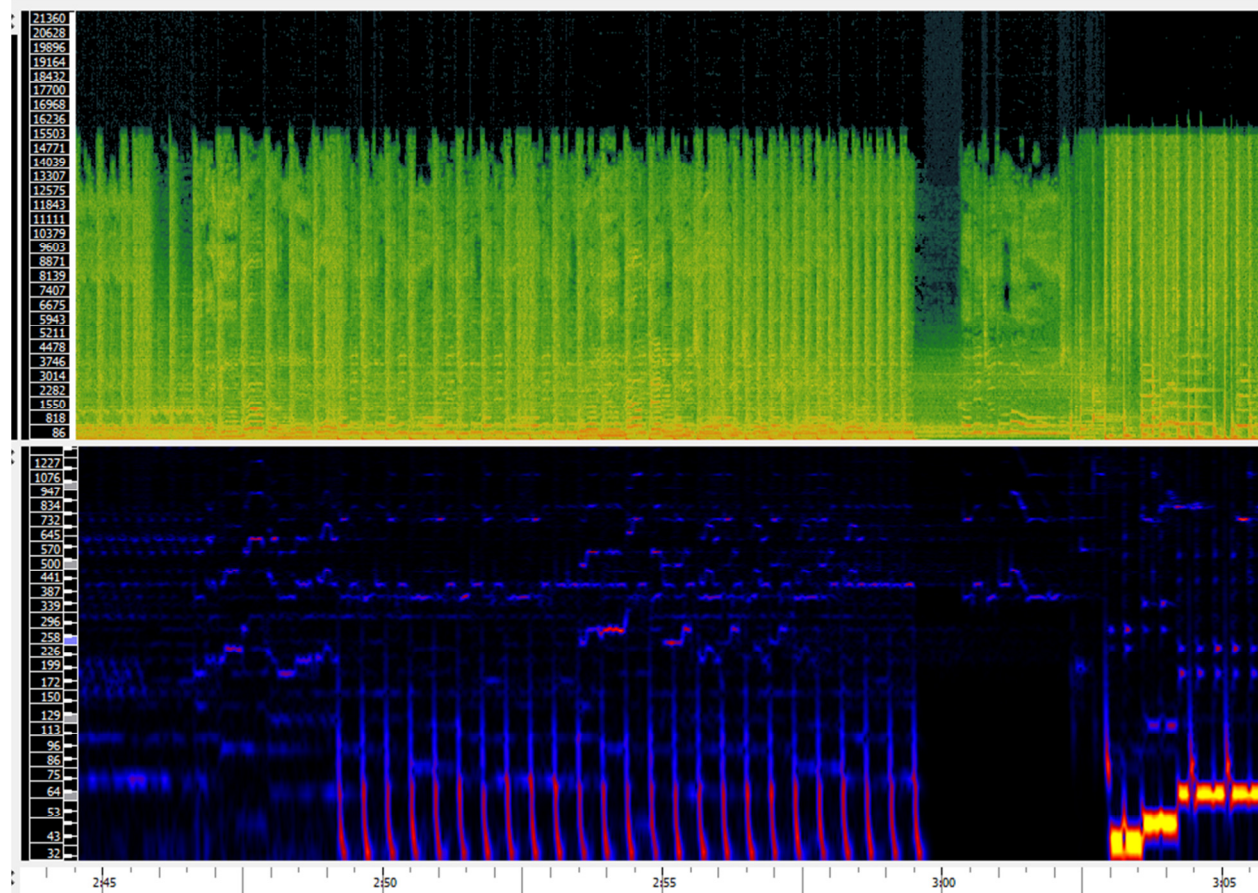
quarter note rhythm (with a slight pickup preceding each eight-beat cycle). At 2:49, it is joined by a kick drum (represented by red “flames” in the melodic range spectrogram in Example 4.12.a) playing a completely-steady quarter note rhythm. This near rhythmic-unison lasts for 16 beats, until the snare-clap accelerates to play eighth notes (2:56), building up more tension that peaks at the break (3:00), and is released at the drop.

a) 2:30-2:50: Adding snare-clap following drum-less verse.



Example 4.12. *Jack Ü, “Where are Ü Now” featuring Justin Bieber (2015), 2:34-3:07.*

b) 2:45-3:05: Adding kick drum (2:49), accelerating snare-clap (2:56), and break (3:00).



Example 4.12. *Cont.*

Summary of Techniques for Increasing Tension

We can see that all of the techniques described in this chapter have related traits. Linear ascending gestures, filter sweeps and stepwise addition, both direct the listener's attention upwards in a goal-oriented manner, in anticipation of a climax. The difference between the two is usually in the duration of the process and degree of intensity. This is also the case with a break at the end of a buildup and a milder removal of sonic layers. Both perform a similar action, but

differ in intensity and structural role. Even rhythmic buildups, as I mentioned, have much in common with sonic density buildups, whether linear or stepwise.

Nearly every contemporary pop song contains a combination of filter sweeps, rhythmic intensification, step-by-step addition of sonic layers, abrupt removal of sonic layers, and breaks. Each of these techniques is strategically used as a building block in a sonic narrative. Stepwise addition is usually effective for slower, prolonged buildups of tension. Short filter sweeps are primarily used to emphasize anticipation toward a new section. Long filter sweeps define entire sections and precede climactic structural arrivals. Abrupt removal of sonic layers significantly heightens tension. Breaks bring the tension to a peak at the end of already-strong anticipatory gestures. Each of these techniques fulfills the buildup function in the setup-buildup-peak sonic cycle, on a different structural level.

RESOLVING/DEFUSING TENSION

I have discussed at length the fact that contemporary pop music rarely exhibits moments in which there is resolution similar to that of the closing tonic. Since a large part of the work I have done on this project has had to do with exploring the ways in which ideas of common practice tonality manifest themselves in the sonic dimension of today's pop, the issue of closure has consistently occupied my thoughts. At first, I saw a striking resemblance between the riser-drop sequence and a dominant-tonic relationship. However, the drop itself usually represents an arrival at a highly charged section, and therefore cannot function as the "sonic tonic." I later tried to imagine the buildup as a predominant and the drop as a dominant, but that did not quite work for me either. A big riser such as the one in "Till the World Ends" or "Heroes" represents a level

of certainty in the outcome that is not usually present in the predominant function. It is true that in certain structural contexts we do feel that the arrival of the dominant is inevitable, but even in those contexts, the presence of the predominant is not nearly as pronounced as that of the riser in pop songs.

Thus, my conclusion was that there is no one-to-one correlation between these sonic functions and traditional harmonic functions, despite the many similarities. None of the functions in the setup-buildup-peak sequence represent resting points. One of the requirements for achieving a closing tonic in common practice tonality is that all of the different elements “agree” that it is closure time. The harmony, melody, and meter, all have to arrive at the same time via a prescribed trajectory. This is difficult to achieve in today’s pop music, in particular because the chord loops, as I discussed in Chapter 3, do not normally lend themselves to cadential closure. In the rare case that they do, such as in the example of Tove Lo’s “Timebomb,” a cadence can indeed be achieved. However, closure like that of the closing tonic is too scarce in this music to motivate such a correlation.

Though it is impossible to have a single sonic function that parallels a closing tonic, there are numerous partial or incomplete resolutions. For example, there are arrivals in which one anticipatory gesture or process is resolved on a local level, but a larger anticipatory process continues on a higher structural level. This frequently occurs in stepwise sonic additions. Each step is connected to the next by a small anticipatory gesture, such as a noise sweep. The noise sweep is resolved each time, but the larger process carries on, as sonic layers continue to be added in anticipation of a climax.

On a larger scale, the moment of arrival at a drop following a riser is, in some ways, more complicated to define. On the one hand, the anticipatory process leading up to the climax is a structurally significant one that is not perceived as a mere step in a larger process. The drop is what the listener has been anticipating throughout the entire process, much like the tonic is anticipated during the unfolding of a phrase. On the other hand, the climactic section triggered by the drop constitutes a highly charged and anticipatory passage, awaiting resolution as well. So, in a sense, there are two structurally significant arrivals: The first, when the climax is reached (the drop), and the second, when the climactic passage ends and the high level of sonic energy is, to a certain extent, defused, though this arrival still does not constitute a complete resolution.

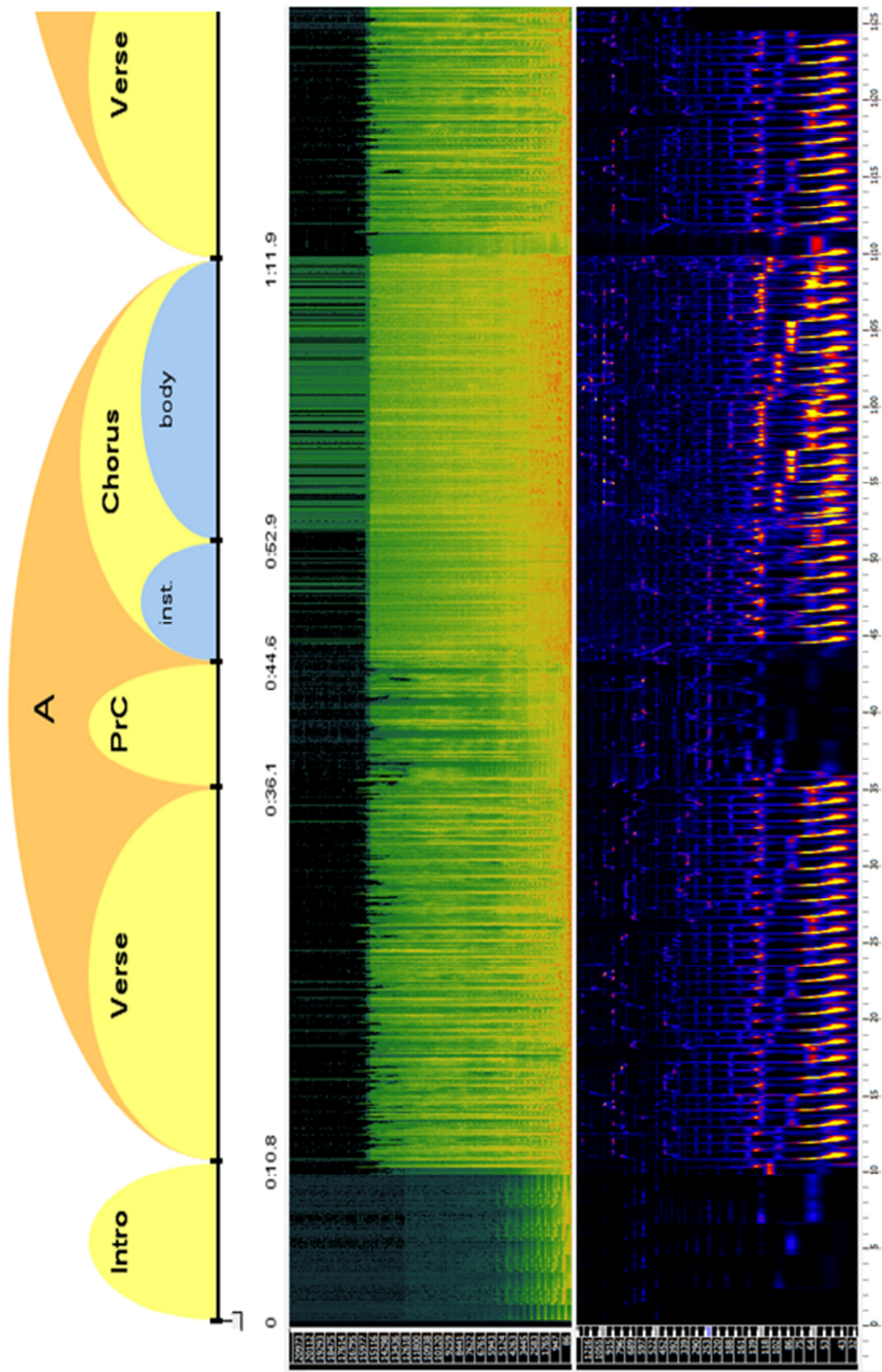
Common Structural Points of Diminishing Tension

In a typical pop song, there are two moments at which the tension is considerably defused – the downbeat of the second verse (usually following the first chorus), and the end of the song. Neither provides complete closure in most songs, as this would require all of the musical elements to participate in achieving this closure, but these are the two moments at which the most significant diminishing of energy occurs.

Chorus-to-Verse

A structural moment at which there is commonly significant diminishing of tension in a typical pop song is the downbeat of the second verse. After building up energy and reaching a sonic peak in the chorus, the energy usually drops to a level relatively close to the song's baseline energy level. This signals the end of the first major section in the song (verse—

pre-chorus—chorus) and provides some relief from the high level of tension in the chorus. We can hear this, for instance, in Demi Lovato’s “Cool for the Summer” (2015). The sonic energy from the chorus (0:44-1:12) drops abruptly in the verse that follows immediately. However, unlike other instances of sudden energetic drops, the overall effect is not to heighten tension, but to temporarily reduce it. We can see in example 4.13 that the sonic energy in the second verse (starting at 1:11) is significantly lower than in the chorus, but somewhat higher than in the first verse (0:10-0:36). The moment of arrival at the second verse creates a balance of sorts between the falling energy that ends the chorus and the continuing process of accumulating energy that began in the first verse. It is not complete closure, as another process with anticipatory properties begins, but it is a moment in which the tension is temporarily scaled back.

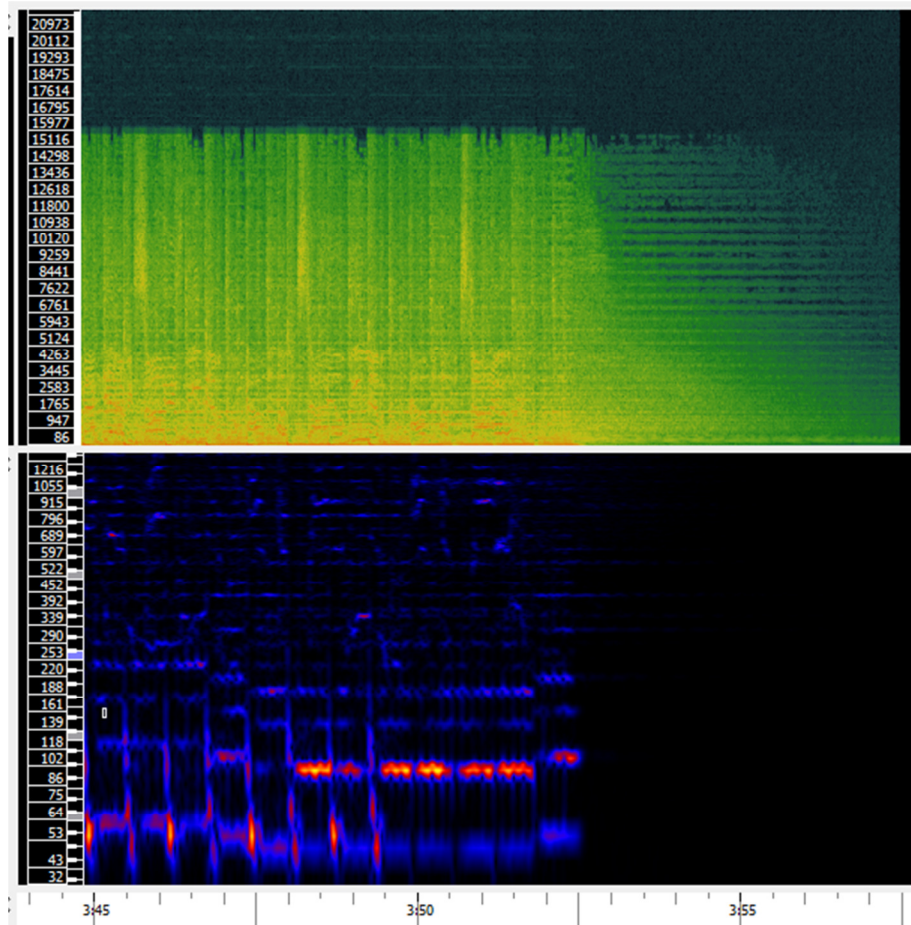


Example 4.13. Demi Lovato, "Cool for the Summer" (2015), 0:00-1:20.

Abrupt Song Endings

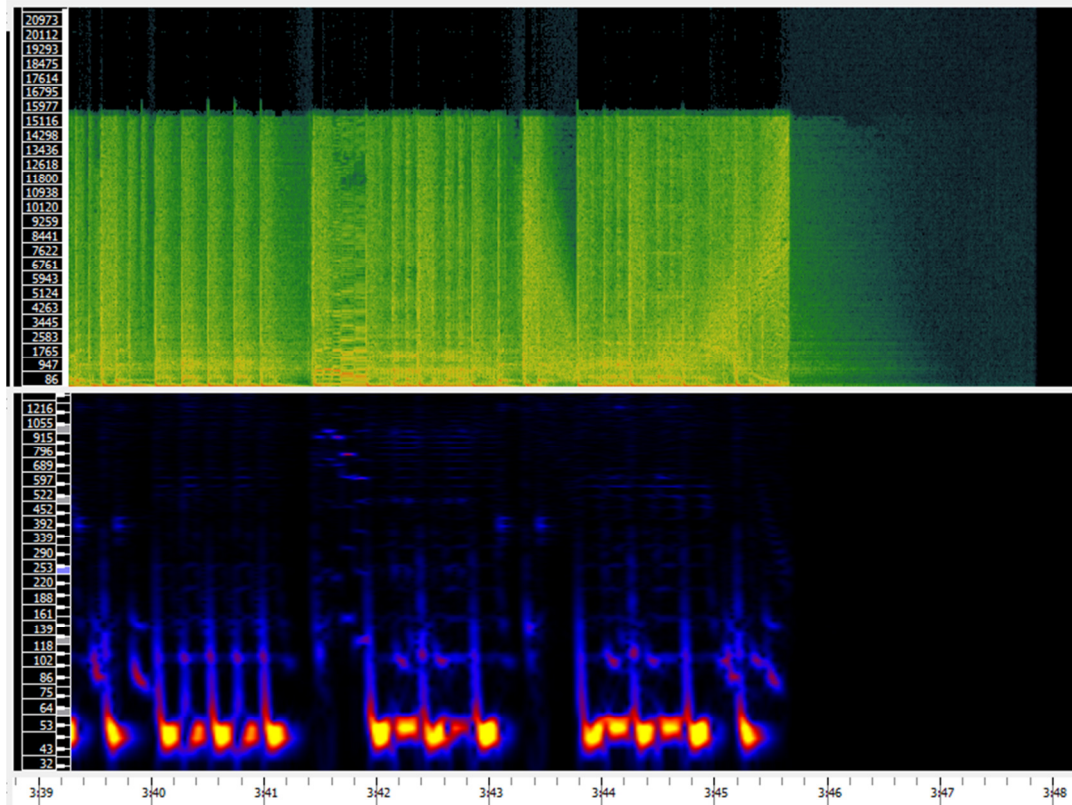
Typical song endings in this genre exemplify the absence of closure in the setup-buildup-peak sequence. Most songs simply conclude at the end of a peak section. There may be a lingering echo or a syncopated vocal trail, but the song generally comes to an abrupt halt. Clearly, this does not constitute closure, nor could it become closure by falling back to a setup, because the setup is also an anticipatory function. In most other genres, this type of ending would represent an extremely tense gesture (e.g., Mozart ending a piece on the dominant). However, this type of ending is so idiomatic and predictable in today's pop that it is akin to an arrival point at an expected outcome, while not setting up anticipation toward the next event. Thus, while there is no complete closure in this situation, the tension is nevertheless diminished.

One example (of many) of this type of ending can be heard in Taylor Swift's "Style" (2014). The bass ends on a strongly syncopated gesture – four repeated sixteenth notes played on the final beat of a measure. The music ends abruptly, and the only lingering sounds are a pale, delayed iteration of the accompanying synth riff, and an echo from a noise track, previously "buried" in the mix but now in the foreground of the almost-empty sonic space, quickly fading out. We can see this unfold in Example 4.14. The activity in the melodic range spectrogram (where the amplitude threshold for color visibility is higher in order to allow individual sounds to be clearly represented) comes to a halt, and the plain spectrogram gradually becomes darker (i.e., all frequencies significantly decrease in amplitude). The plain spectrogram shows a descending filter sweep (3:53-3:57), but it is very short and is applied only to the lingering noise, and unlike longer descending sweeps, it does not defuse tension but simply causes the noise to sound as if it is falling down in the sonic space.



Example 4.14. *Taylor Swift, “Style” (2014), 3:45-3:57.*

An even sharper ending occurs in Justin Bieber’s “Beauty and a Beat” featuring Nicki Minaj (2012). We can see a very short (two-second long) lingering echo in the spectrogram (example 4.15), but it is barely audible, and constitutes a very subtle sonic smoothing mechanism. The music effectively ends with a bass attack on beat 3 of the final measure, and a falling noise on beat four, which ends abruptly just before the arrival of the downbeat, creating a highly charged ending.

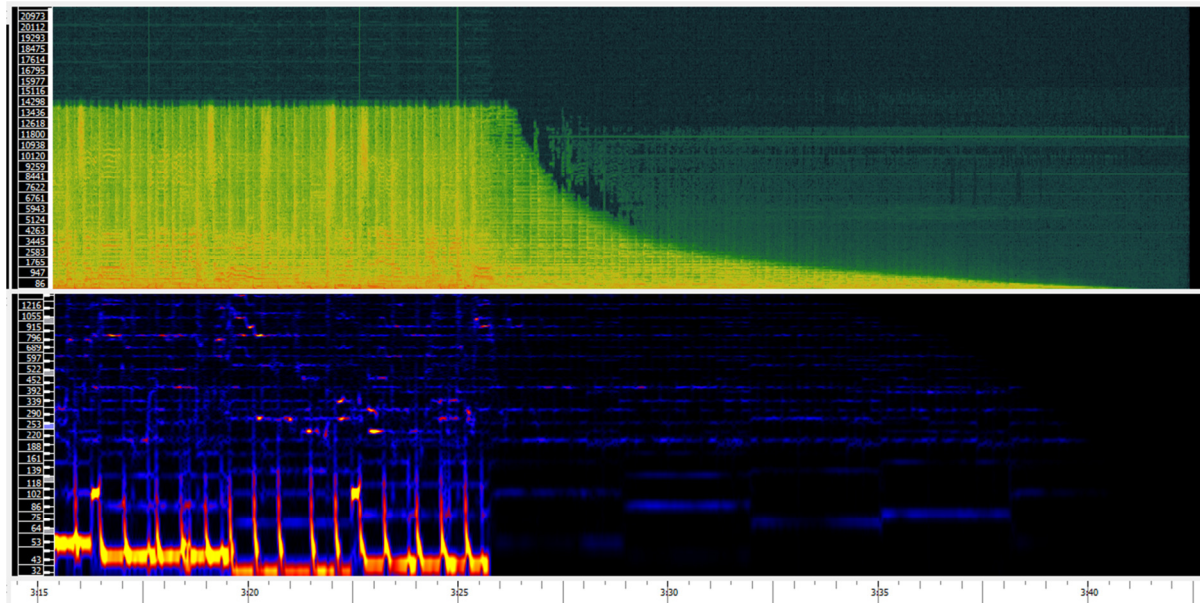


Example 4.15. Justin Bieber featuring Nicki Minaj, “Beauty and a Beat” (2012), 3:39-3:48.

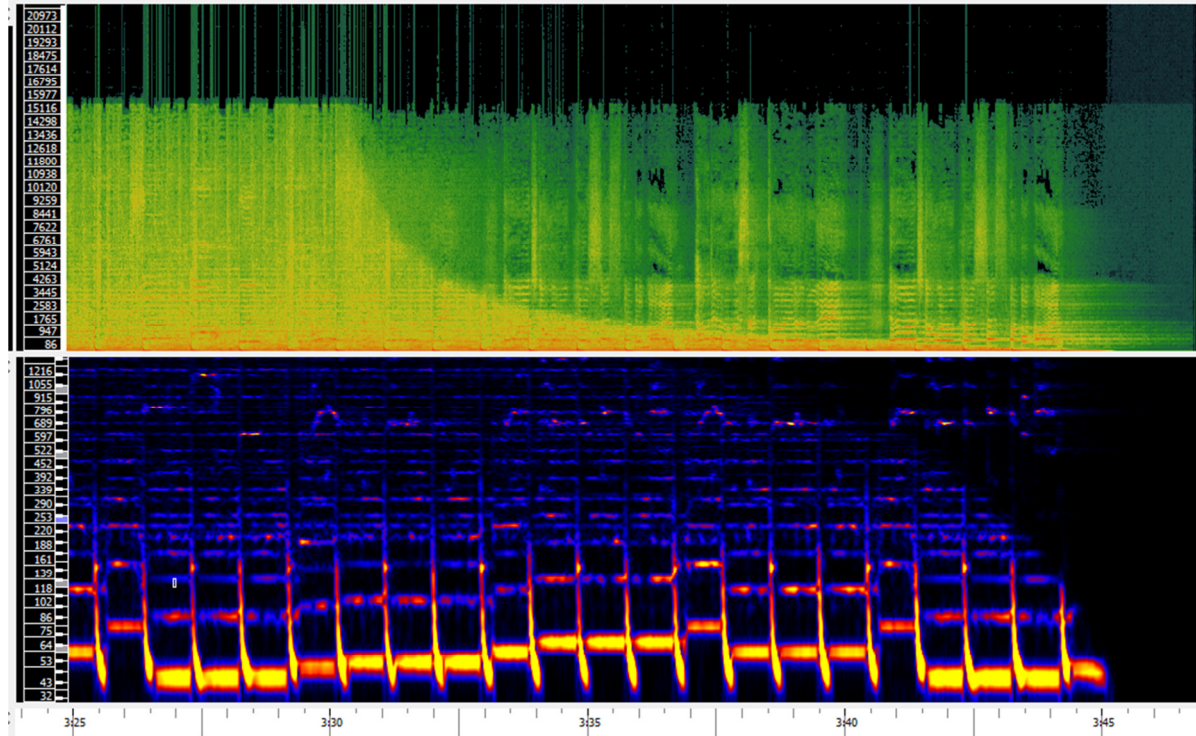
To summarize, there are three primary factors that influence the feeling of tension in endings such as those in “Style” and “Beauty and a Beat.” The first is their abrupt and syncopated character, which increases tension rather than decreasing it. The second is the predictability of these endings due to their prevalence in the genre, which makes them a target of anticipation rather than creators of it. The third is that these endings constitute the first time since the beginning of the song when there is no feeling of anticipation toward an imminent arrival. The last two, in my experience as a listener, counter the first and cause a relative reduction in the level of tension.

Descending Filter Sweep Endings

The descending filter sweep is simply a reversal of the process created in ascending filter sweeps such as those present in risers. The filter is moved to gradually cut high frequencies rather than amplify them. The descending filter sweep is a rare gesture that defies the setup-buildup-peak cycle. Out of the 100 songs that I have analyzed for this dissertation, this occurs in only two songs – Becky G’s “Lovin’ So Hard” (2015) and Nicki Minaj’s aforementioned “The Night is Still Young” (see Examples 4.16 and 4.17). Rather than set up a sonic ascent or build up to a climax, this sweep gradually reduces the tension by “sliding down” from the highly charged final peak section (usually the third chorus or post-chorus).

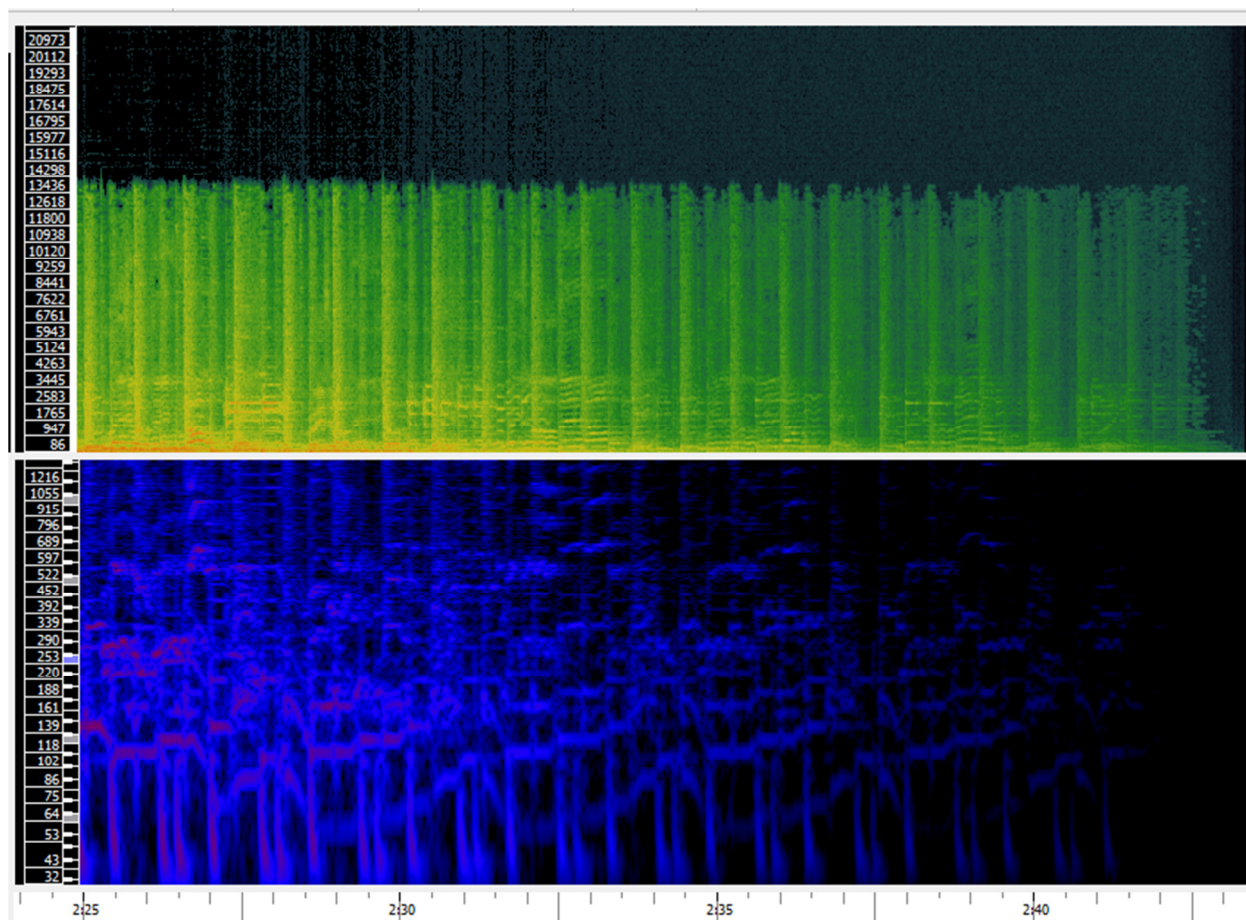


Example 4.16. *Becky G, “Lovin’ So Hard” (3:15-3:42), descending filter sweep.*



Example 4.17. Nicki Minaj, “The Night is Still Young” (2014), 3:25-3:47. *descending filter sweep.*

Descending filter sweeps resemble in their function the popular fadeout ending in 20th-century rock and pop music, in which the final chorus is repeated while the master volume control fades the volume. The main difference is that the reverse filter sweep is a timbral process, in which the gradual cutting of high frequencies causes the sound to become darker over time. In fadeouts, there is no discrimination between frequency bands, so the timbre remains intact while the volume fades out. We can see a visualization of a fadeout at the end of Queen’s “Crazy Little Thing Called Love” (1980) in Example 4.18 – the amplitude of all frequencies fades simultaneously, with hardly any relative difference in the pace of the fade between one frequency band and another.

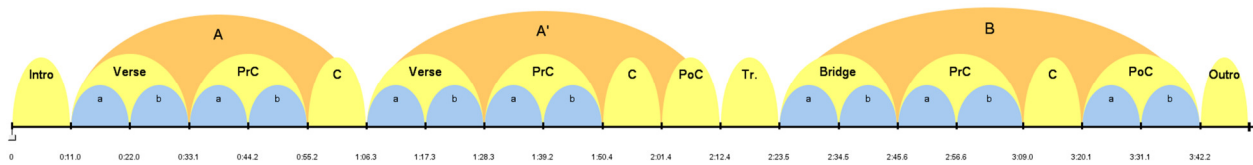


Example 4.18. Queen, “Crazy Little Thing Called Love” (1980), 2:25-2:43. Fadeout.

THE BIG PICTURE

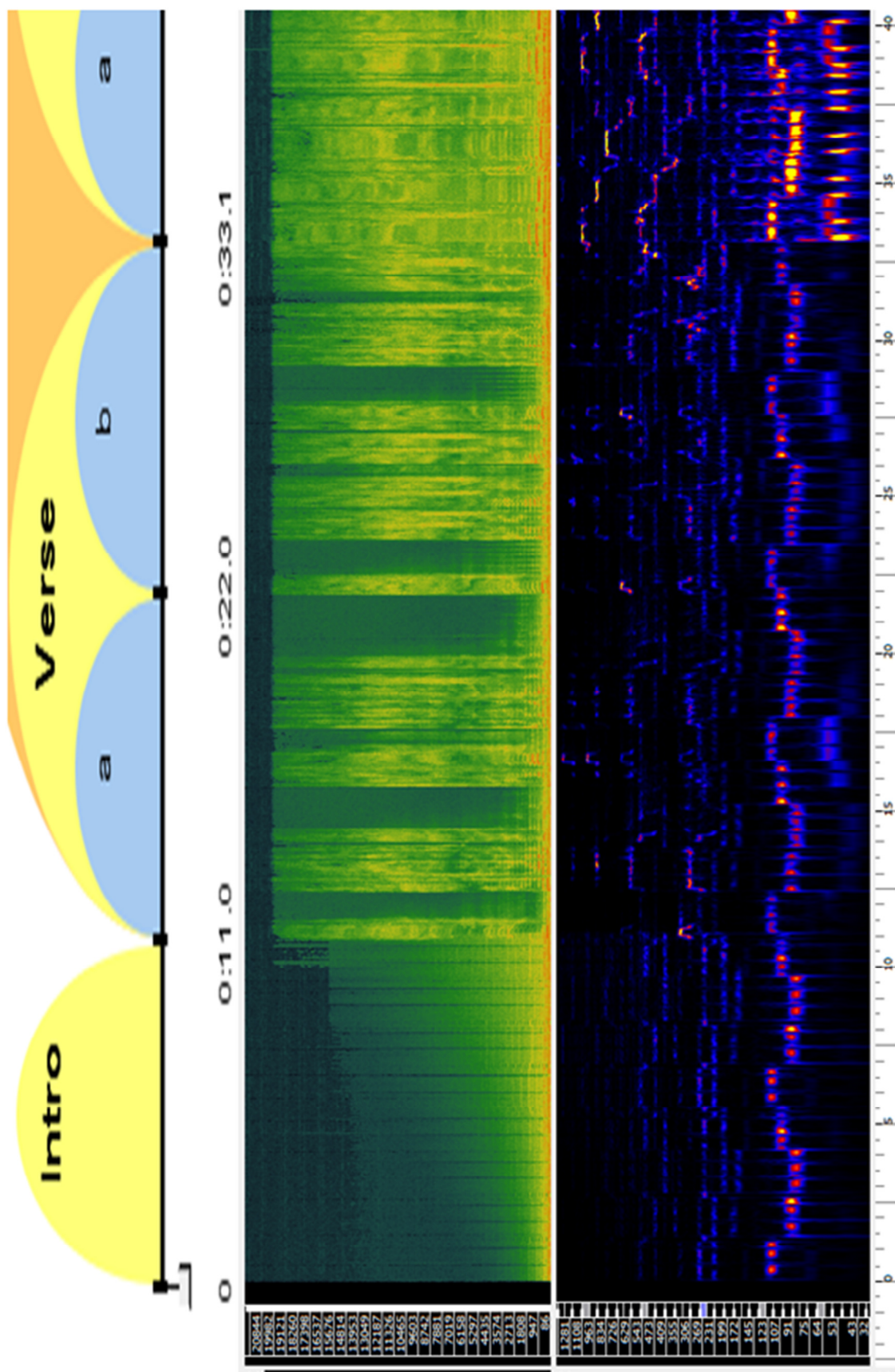
I have laid out the various individual ways in which tension is commonly increased and released in today’s pop songs. To conclude this chapter, I will demonstrate how these individual gestures combine to help shape a sonic narrative in a song. The song I will use for this purpose is Ellie Goulding’s “Burn” (2012). This song is in an **AA’B** form, in which **A** (0:11-1:06) includes a verse, pre-chorus, and chorus, **A’** (1:06-2:12) repeats them and adds a short post-chorus, and **B** (2:23-3:42) consists of a bridge, pre-chorus, chorus, and a longer post-chorus. For the sake of

clarity, I will refer to **A**, **A'**, and **B** as “large sections,” in contrast with internal sections such as verses and choruses, to which I will simply refer as “sections.” As we can see in Example 4.19, each large section can be divided into 11-second-long blocks of music (each spanning the duration of two chord loops). Some sections (e.g., verse, chorus, etc.) span two blocks and others only one. Each block has a local anticipatory role, while at the same time participating in longer-span anticipatory processes.

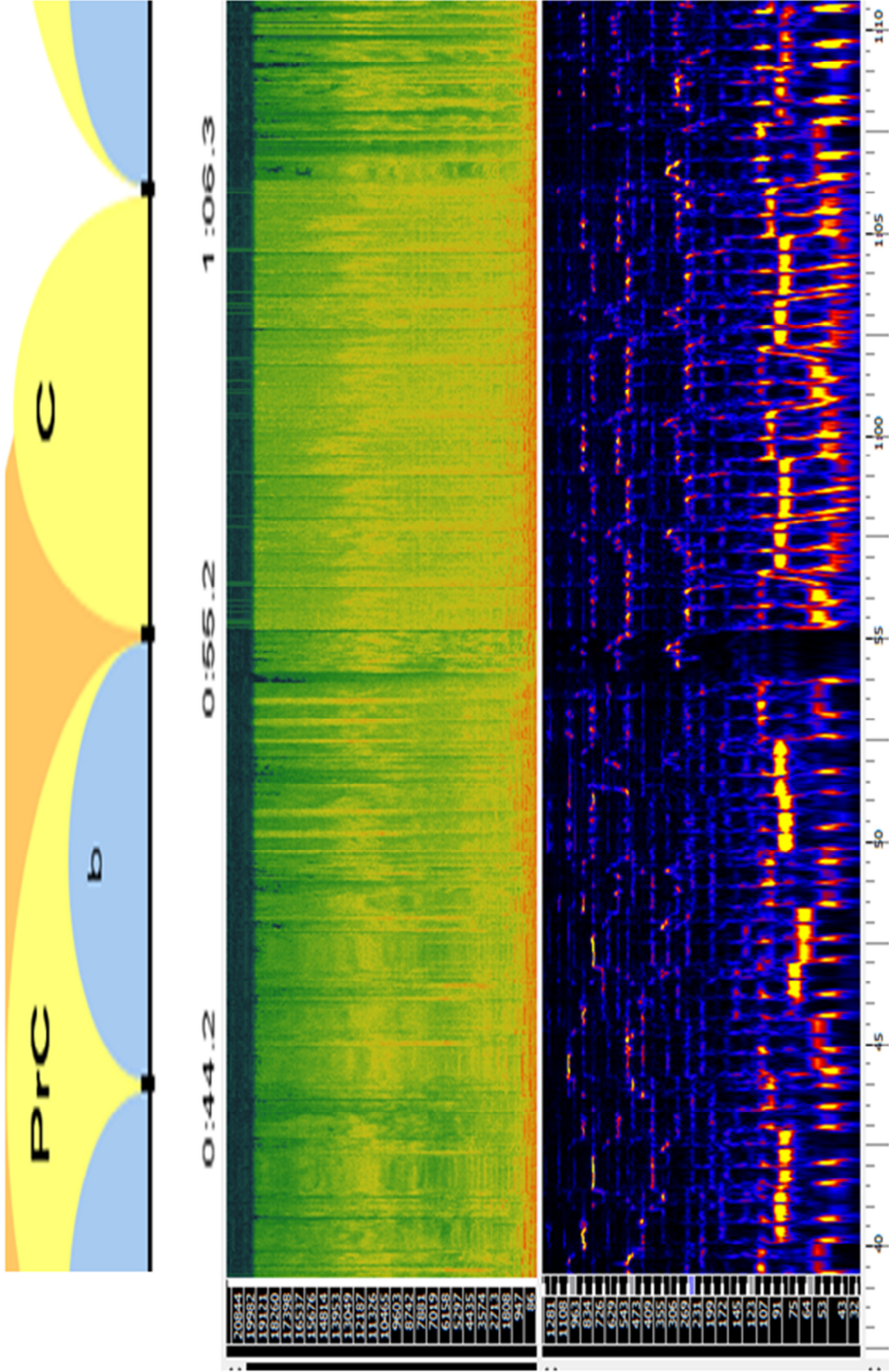


Example 4.19. *Ellie Goulding, “Burn” (2012).*

Glancing at the spectrogram of the intro and **A** (0:00-1:06) in example 4.20, the general process is quite clear. The intro leads up to the verse via a filter sweep, and from this point the sonic density increases at the beginning of every block (with a significant increase occurring at 0:33, which correspond to the beginning of the pre-chorus), until it climaxes at the chorus (0:55-1:06). However, this does not tell the whole story. Following the intro, the verse begins (0:11) with Goulding’s voice accompanied by a soft, heavily reverbed guitar sound, along with a barely-audible rhythmic plucking sound in the far background. This plucking sound moves to the accompaniment’s foreground in the second half of the verse (0:22), bringing the sonic energy a step forward. The arrival of the pre-chorus (0:33) is emphasized with a drum hit, which triggers a white noise and a bass sound. At the second chord loop, arpeggiation is added in the high register of the synth, Goulding’s vocal line is doubled an octave below, and she is harmonized by backup vocals. Finally, chimes enter in mid-loop at 0:50.



Example 4.20. *Ellie Goulding, "Burn," 0:00-1:10.*

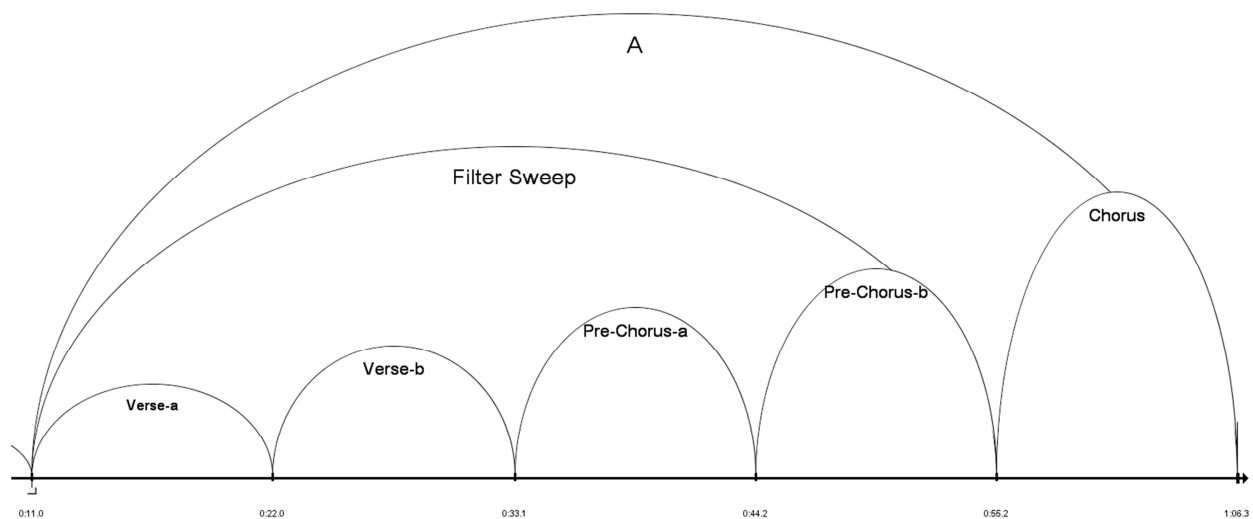


Example 4.20. *Cont.*

So far, this sounds like a rather ordinary step-by-step addition of sonic energy. However, if while listening to the recording of the song, it feels like some of the sounds are sneaking up on you, it is because on top of the stepwise addition, there is a subtle filter-sweep throughout the entire process, gradually brightening the mix and moving sounds from the background to the foreground. When this happens, it causes us to feel like we are just noticing sounds that have been there the entire time. While it is not the type of overt filter sweep that leads the listener toward the climax, it nevertheless affects the tension buildup process by adding a linear ascending dimension to it. Because it is so subtle, it is easy to miss, both when listening to the song and looking at the spectrogram, but even if we are not fully conscious of this effect, it still influences our listening experience. This is an instance in which the close listening mode I described in chapter 2 is helpful in noticing the more covert driving forces behind a song.

Example 4.21 offers a rough visualization of the development of sonic energy in **A** – a layer of sonic density is added at the beginning of each subsection, making it a stepwise buildup, but the ascending filter sweep occurring throughout the large section adds an overarching linear element to it. The process in **A'** is extremely similar. However, it begins on a much higher baseline energy level. That is, the second verse (1:06-1:28) is far denser and rhythmically active than the first verse (0:11-0:33). In fact, if we listen to the first pre-chorus (0:33-0:55) and skip directly to the second verse (omitting the chorus), we can hear that the second verse continues more or less where the pre-chorus left off, from a sonic energy perspective. The climax (i.e., the second chorus, 1:50-2:01), on the other hand, is identical to the first chorus (0:55-1:06). This means that overall, **A'** is more intense than **A**, but the energetic gap between each step in the buildup is smaller. Additionally, because **A'** has a relatively substantial presence of high frequencies to begin with, the overarching filter sweep is even less noticeable than it is in **A**.

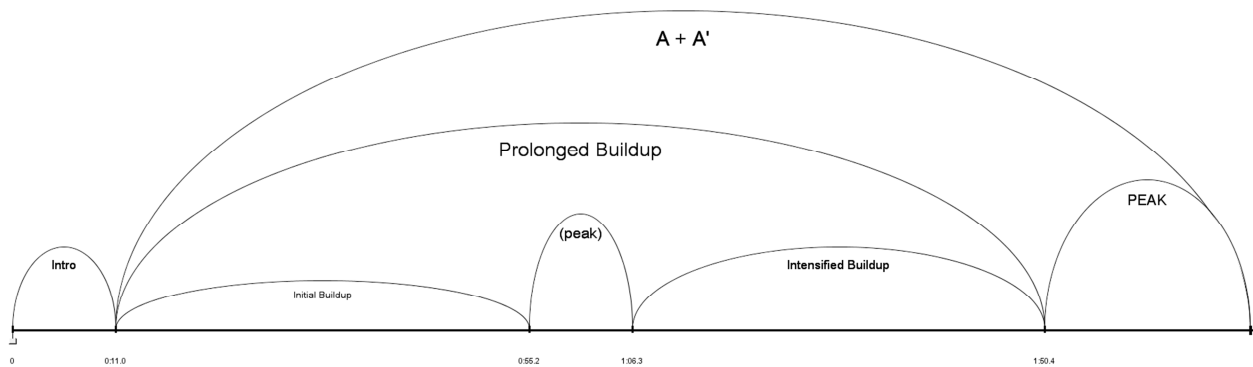
However, it is still there, gradually brightening the overall sound of the mix. This can be heard more clearly by skipping back and forth between the first half of the second verse (1:06-1:17) and the second half of the second pre-chorus (1:39-1:50). The latter is much brighter in its overall sound.



Example 4.21. Rough visualization of the sonic energy growth in Ellie Goulding's "Burn," 0:11-1:06.

While the first and second choruses are identical, and therefore **A** and **A'** arrive at an identical climax, the second chorus is prolonged by a post-chorus, in which the higher frequency range becomes denser due to an arpeggiating high synth along with Goulding singing in her high register. So, while the point of arrival is identical for both **A** and **A'**, the sustained high-energy section that follows the point of arrival in **A'** is twice as long as the one in **A**, and ultimately reaches a higher peak in its second half. If we take a pseudo-Schenkerian approach and try to capture the higher-level process that takes place in the first two large sections of this song, we

can interpret it as illustrated in Example 4.22, where everything prior to the second chorus is a continuous buildup, with a relatively short interruption in the middle, in the form of the first chorus. The primary reasons for this interpretation are that: a) the buildup in **A'** starts almost exactly, sonically speaking, where the buildup in **A** ended; and b) the peak in **A'** is sustained for a longer period of time and reaches a higher level of sonic energy, despite the initial arrival being identical to the arrival in **A**. Note that I have not included a setup function in this illustration, since the overarching filter sweep transforms the stepwise addition of sonic layers into a smoother continuous process, rather than a setup and a buildup.

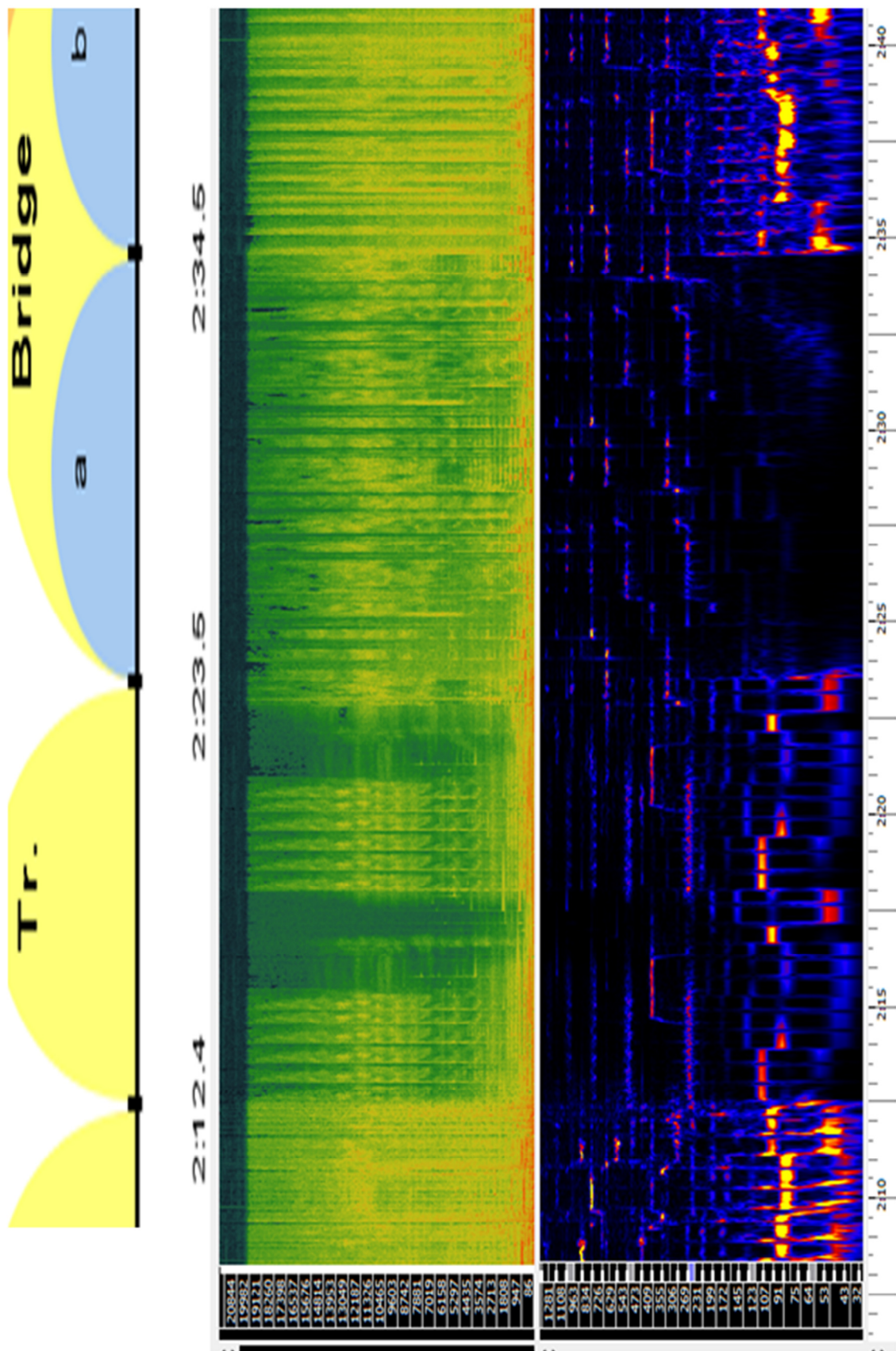


Example 4.22. *Reduction of sonic development in Ellie Goulding, “Burn,” 0:11-2:12.*

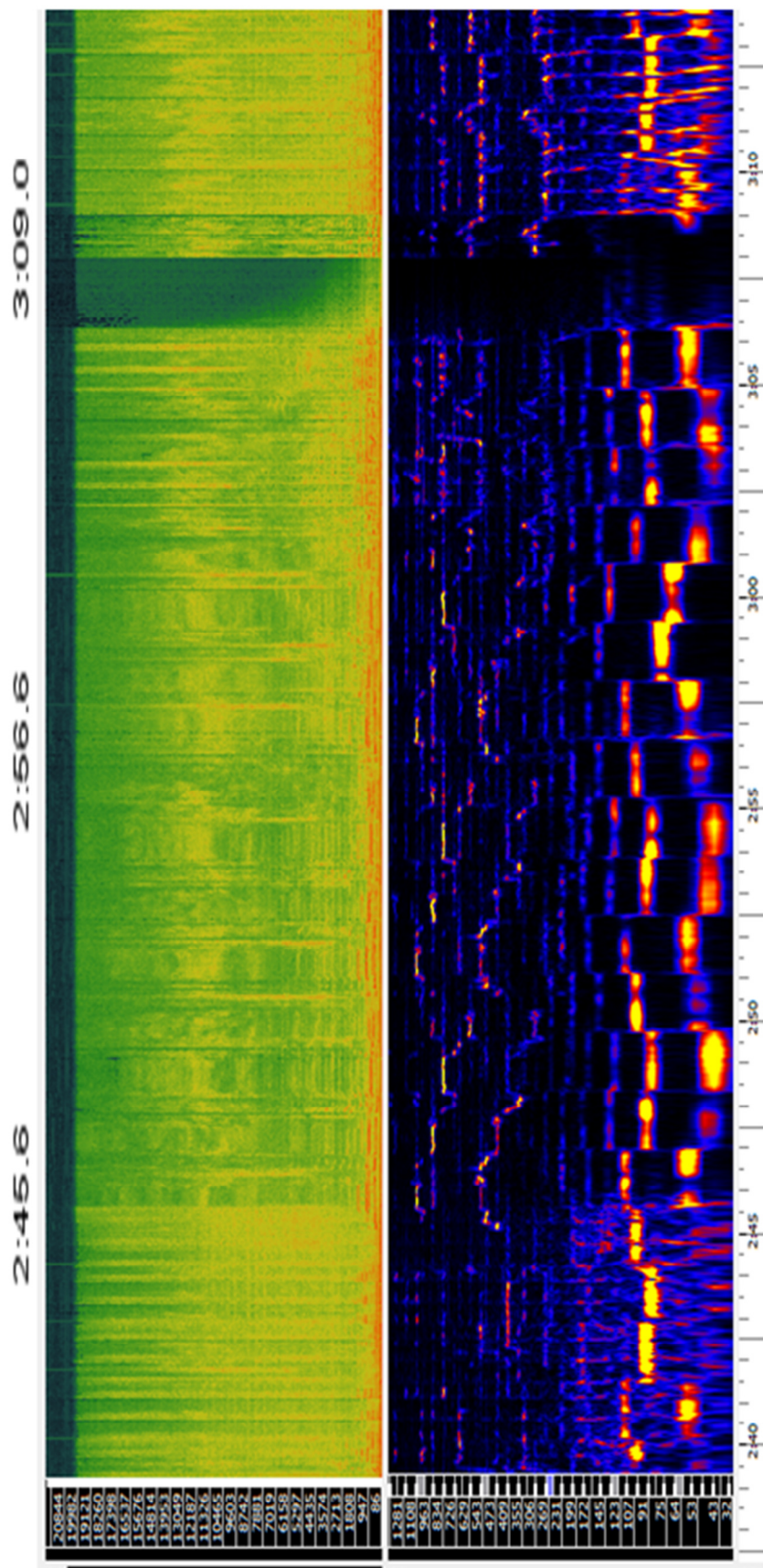
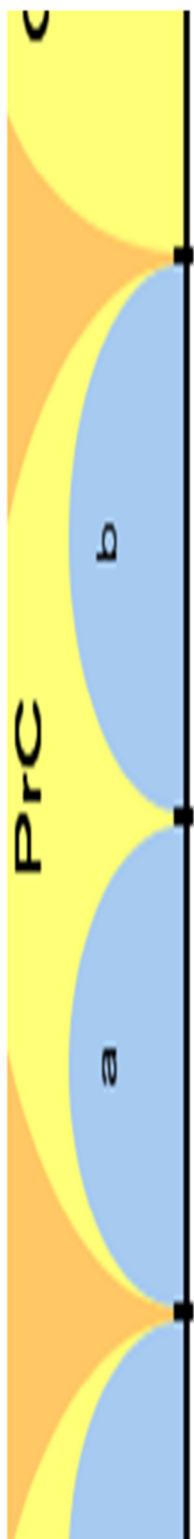
While up to the end of **A'** we see a relatively slow and continuous buildup, the road to the climax in **B** is faster, choppy, and more intense. **B** starts lower than **A** in terms of sonic energy, and ends higher, extending the range and the intensity of the buildup. In addition, the sustained climactic section (chorus + prolonged post-chorus) is not only longer than the previous ones, but it is also higher in sonic energy.

After **A'** ends, we hear a series of sonic falls (i.e., abrupt dropping of sonic layers), which lower the sonic energy but increase the tension (see Example 4.23). The first sonic fall occurs at the transition (2:12-2:23), which would likely function as the beginning of the bridge, if not for the additional fall at 2:23, as seen in the melodic range spectrogram. The full texture at the post-chorus suddenly empties out, leaving only the bass, the (heavily reverbed) vocals, and otherwise very sparse accompaniment. The transition ends with a small noise sweep, deceiving us into expecting an arrival at a sonically denser section, but instead the sonic energy falls even further, as the bass drops out, leaving the vocals to be accompanied by soft piano chords.

B begins only now, with the arrival at the bridge (2:23), which is normally the lowest point of sonic energy in a contemporary pop song. The second half of the bridge adds a progressively intensifying drum beat, again causing us to expect a climax at the hypermetric downbeat. Instead of a climax, we get another fall (2:46), this time to an additional iteration of the pre-chorus, but without the drums and the synth riff. However, this fall does not drop to the energy level of the beginning of the bridge. Though the drums drop out, the music retains its rhythmic character via an arpeggiated synth accompaniment, intensified by an ascending noise and a subtle filter sweep similar to the one in **A** and **A'**. A prolonged break signals the arrival of the chorus, which, as I mentioned, is sonically more intense than its previous iterations. The chorus is followed by a prolonged post-chorus, which further intensifies the climax, particularly in its last part, where for the first time we hear a contrapuntal vocal line competing with the lead vocals.

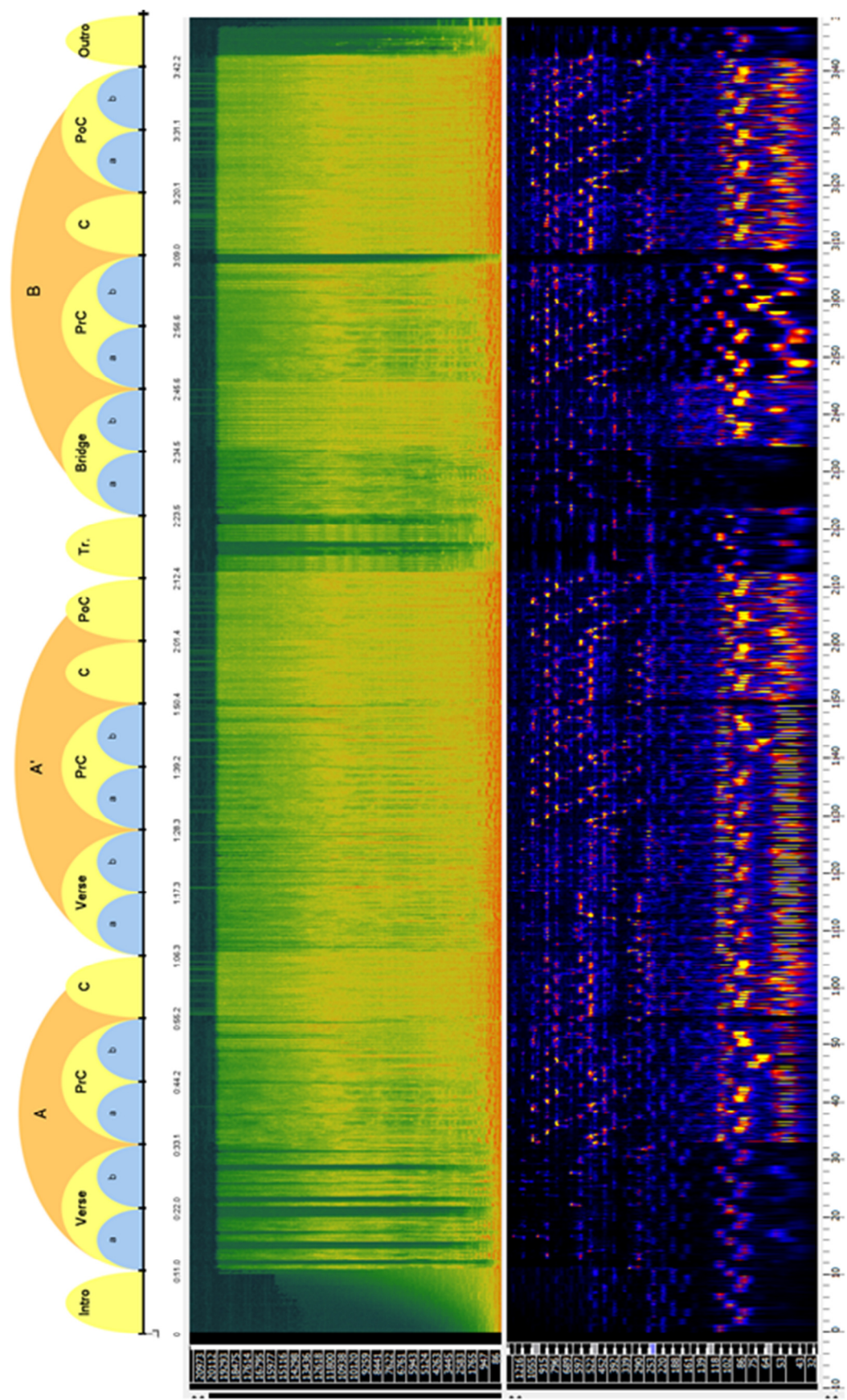


Example 4.23. *Ellie Goulding, "Burn," 2:09-3:15.*



Example 4.23. *Cont.*

The buildup process in **B** can be interpreted as a more intense version of the aggregate buildup in **A** through **A'**. Again, we have two separate buildups (in the bridge and the pre-chorus) that combine to forge a larger buildup, with an interruption in the middle. The bridge begins at a lower point than the first verse, and **B** ends on a higher climax than **A'**. Therefore, the buildup in **B** starts with less sonic energy but more tension, which is built upon by covering more sonic ground in a shorter timeframe. Going back to the beginning of the song, the intro is characterized by a miniature buildup in the form of a filter sweep applied to the synth riff, and this is mirrored by the outro as well. So, in fact, this song has two large-scale structural buildups – one long and relatively moderately-paced (**A—A'**), and one shorter, quicker, and more intense (**B**) – and two symbolic, miniature buildups in the intro and outro. This is visually supported by a zoomed-out view of the spectrogram in example 4.24. We can see three areas of peak energy – 0:55-1:06, 1:50-2:12, and 3:09-3:42. Each new peak is longer than its predecessor. The buildup up to the second peak looks continuous, despite the presence of the first peak, while there is a clear break in the texture following the second peak, which begins an independent and more intense buildup toward the third peak. This entire structure is garnished by the filter sweeps on both ends.



Example 4.24. *Ellie Goulding, “Burn.”*

Another noteworthy occurrence in “Burn” exemplifies the symbiosis between the traditional dimension of today’s pop songs and their sonic dimension – the pre-chorus and post-chorus, as traditionally defined by lyrics, harmony, and melody, are completely identical in this song. What differentiates them is their sonic function – the pre-chorus participates in a buildup, while the post-chorus is an extension of the peak. If not for these sonic developments (for example, if I were to analyze an “unplugged” version of this song, with only piano/guitar and vocals), it would be completely plausible, or even necessary, to interpret the pre-chorus—chorus—post-chorus sequences as one section in **aab** (0:33-1:06) or **aaba** (1:28-2:12) form. The differences in sonic energy and functionality, however, make it clear that these sections represent different stages in the sonic narrative.

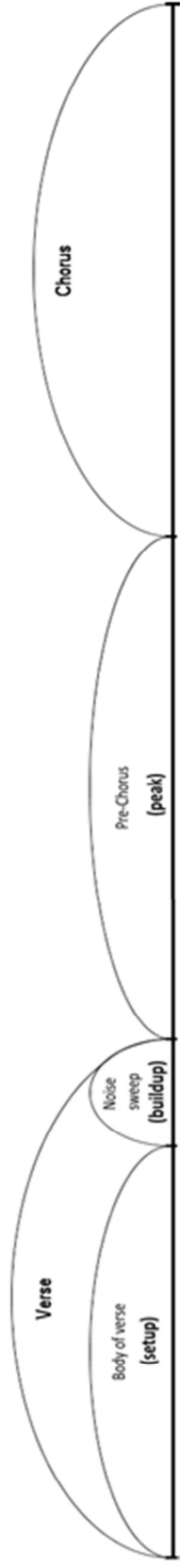
CONCLUSION

This chapter focuses primarily on techniques for building up tension. In Chapter 5, I will explore more in-depth how these techniques function in the context of the setup-buildup-peak sonic cycle. The buildup function is particularly important, because it defines the two other functions on multiple structural levels. For instance, let us take a situation in which the verse functions as a setup, the pre-chorus as a buildup (via stepwise addition of sonic layers), and the chorus as a peak, as illustrated in Example 4.25.a. A short noise sweep leading up to the chorus acts as a lower level buildup. Thus, though on one level the verse is a setup and the pre-chorus is a buildup, on a more surface level, the portion of the verse prior to the noise sweep remains the setup, the noise sweep is a buildup, and the pre-chorus constitutes a peak, as illustrated in Example 4.25.b.

a) Section-level sonic progression.



b) Internal sonic progression between verse and pre-chorus.



Example 4.25 . Sonic functions on two structural levels.

These multi-layered structural relationships constitute perhaps the most striking similarity between pop's sonic syntax and traditional tonal syntax. While, as I previously mentioned, exact parallels cannot be drawn between the sonic cycle and the tonal harmonic cycle, the ability of the sonic syntax to operate on multiple structural levels remarkably resembles traditional tonality. This is extensively exhibited in the song analyses presented in Chapter 5, particularly in the final analysis of Rihanna's "We Found Love," which features an extraordinarily complex sonic structure.

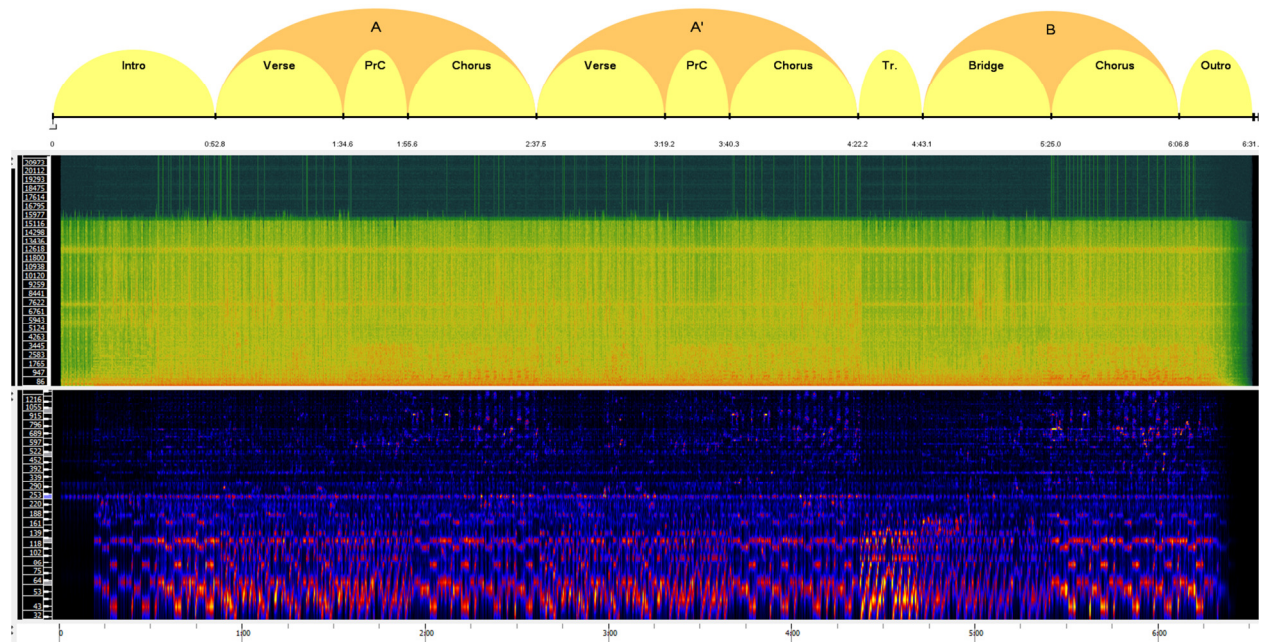
CHAPTER 5

VERSE-CHORUS SONG FORM, AS DEFINED BY MODERN SOUND PRODUCTION

Form in today's pop songs usually operates in two parallel dimensions. The first is the traditional dimension, in which melody and lyrics (and at times, harmony) play a leading role. The second is the sonic density/energy dimension (to which I will henceforth refer as the sonic dimension), wherein timbre and spatialization lead the listener through a song's sonic narrative. This has been true to a certain extent since the dawn of studio-produced popular music. However, the degree to which the latter dimension has become dominant in delineating the form of pop songs in recent years, while they still retain the traditional aspects (unlike pure EDM, for example), places contemporary pop music in a category of its own. Consider the spectrograms visualizing 20th-century pop and rock songs in Example 5.1: two songs revisited from chapter 3, George Michael's "Freedom '90" (1990) and Guns N' Roses's "Sweet Child O' Mine" (1987), as well as Queen's "Crazy Little Thing Called Love" (1980). Although it is possible to detect some timbral shifts in these spectrograms, they are quite subtle. It would be difficult to identify the components of the form simply by looking at the spectrograms, without the aid of the form timelines positioned above them. The only sections that exhibit shifts in sonic density significant

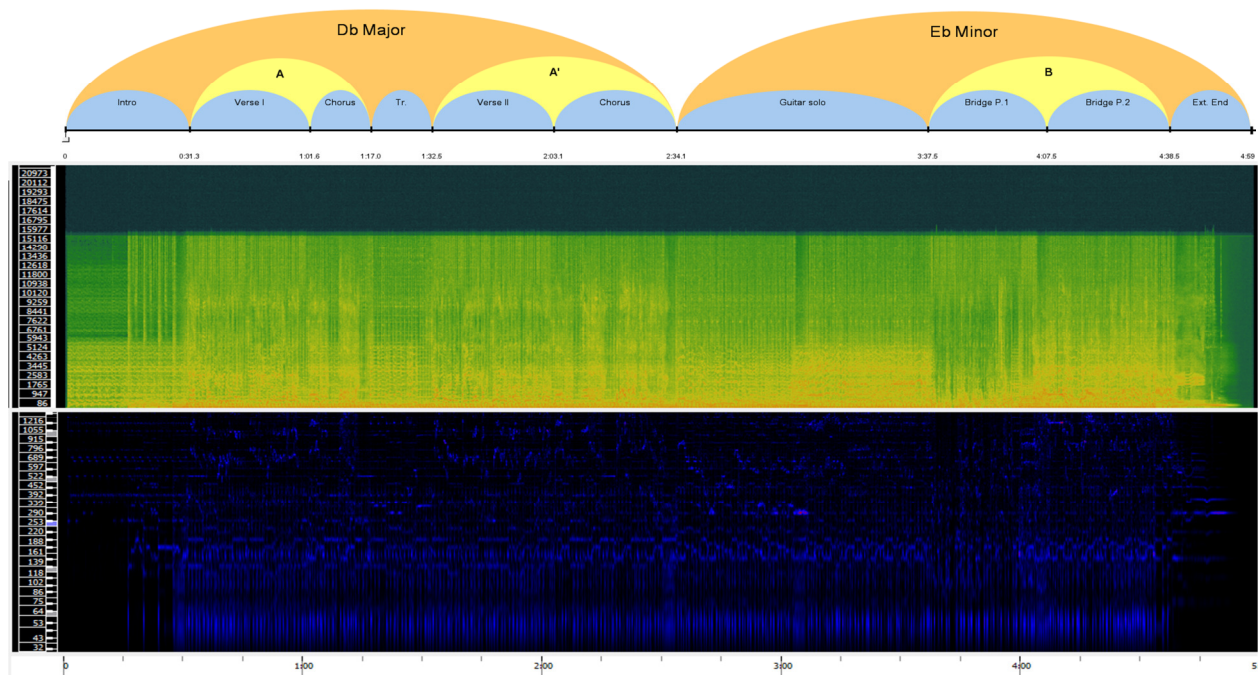
enough to stand out are the intros, outros, transitions, and bridges, where there is a relatively significant drop in energy compared to the other sections.

a) George Michael, “Freedom ‘90” (1990).

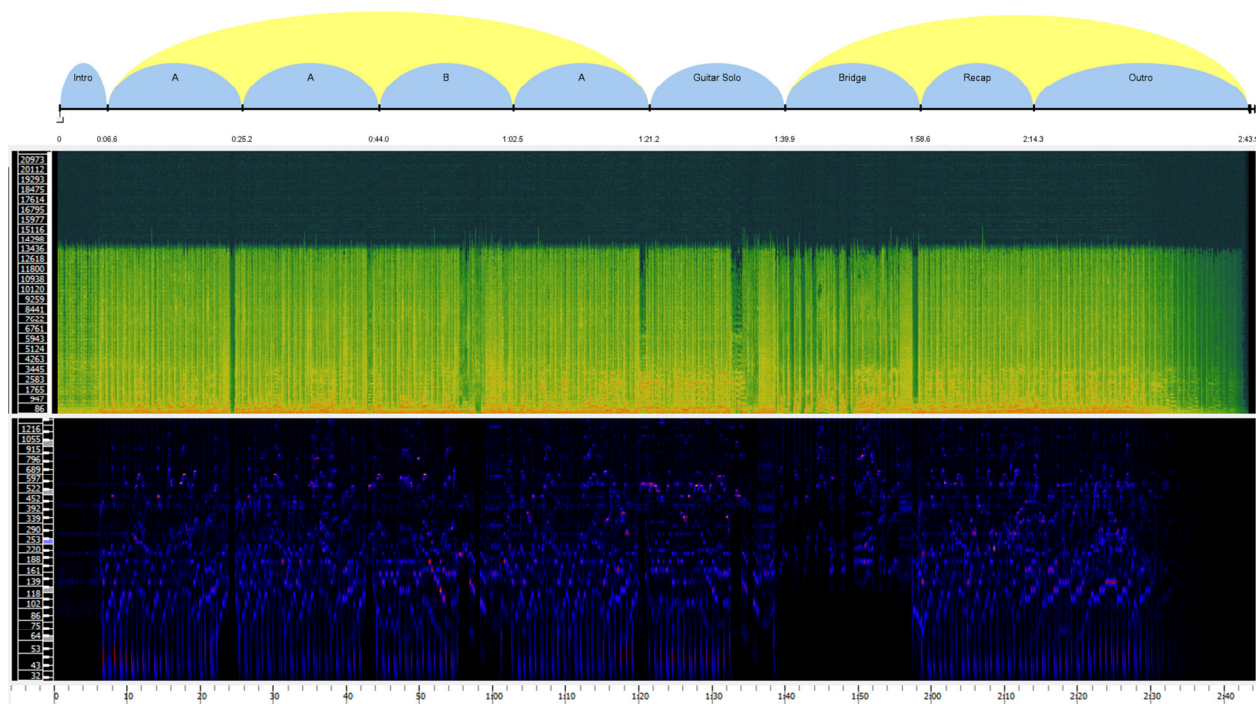


Example 5.1. Full spectrograms and form timelines of 20th-century pop and rock songs.

b) Guns N' Roses, “Sweet Child O’ Mine” (1987).



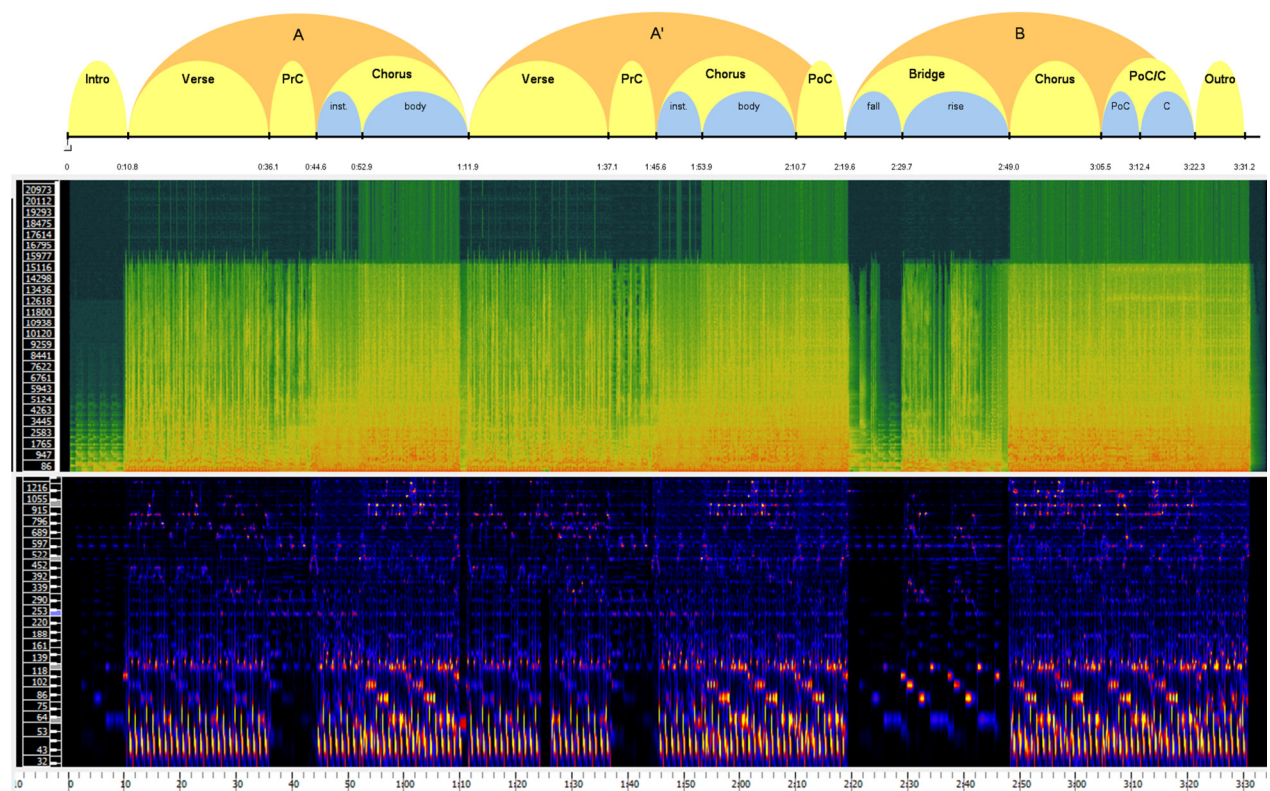
c) Queen, “Crazy Little Thing Called Love” (1980).



Example 5.1. *Cont.*

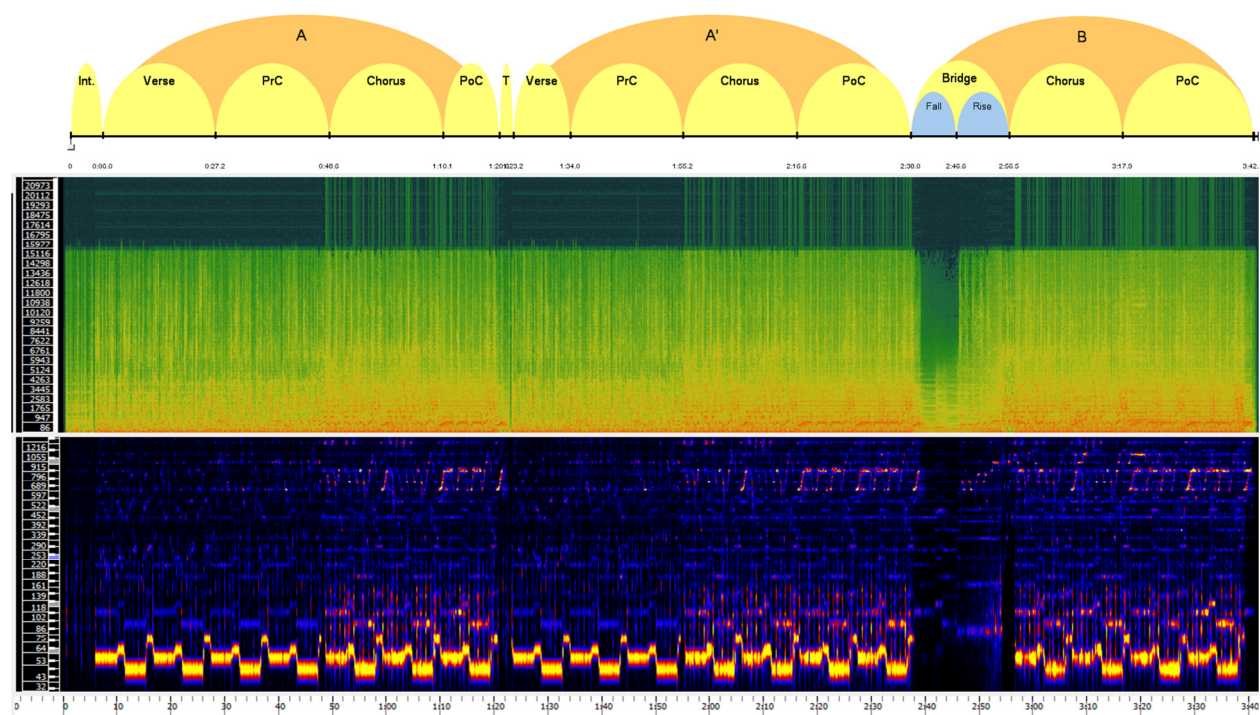
Compare the above spectrograms with those of recent pop songs, shown in Example 5.2: Demi Lovato’s “Cool for the Summer” (2015), Katy Perry’s “Roar” (2013), and Tove Lo’s “Talking Body” (2014). Detecting the sectional boundaries of the form in these spectrograms is a much easier task, due to the substantial differences in sonic density between adjacent sections. Each of these songs features three iterations of the chorus (sometimes with an addition of a post-chorus), represented in the brightest areas of the plain spectrograms. The pre-chorus is also quite clearly visible in “Cool for the Summer” and “Talking Body” as a transition between the verse and chorus. It is less visible in “Roar,” because the sonic development between the verse and the chorus in this song is relatively minimal.

a) Demi Lovato, “Cool for the Summer” (2015).

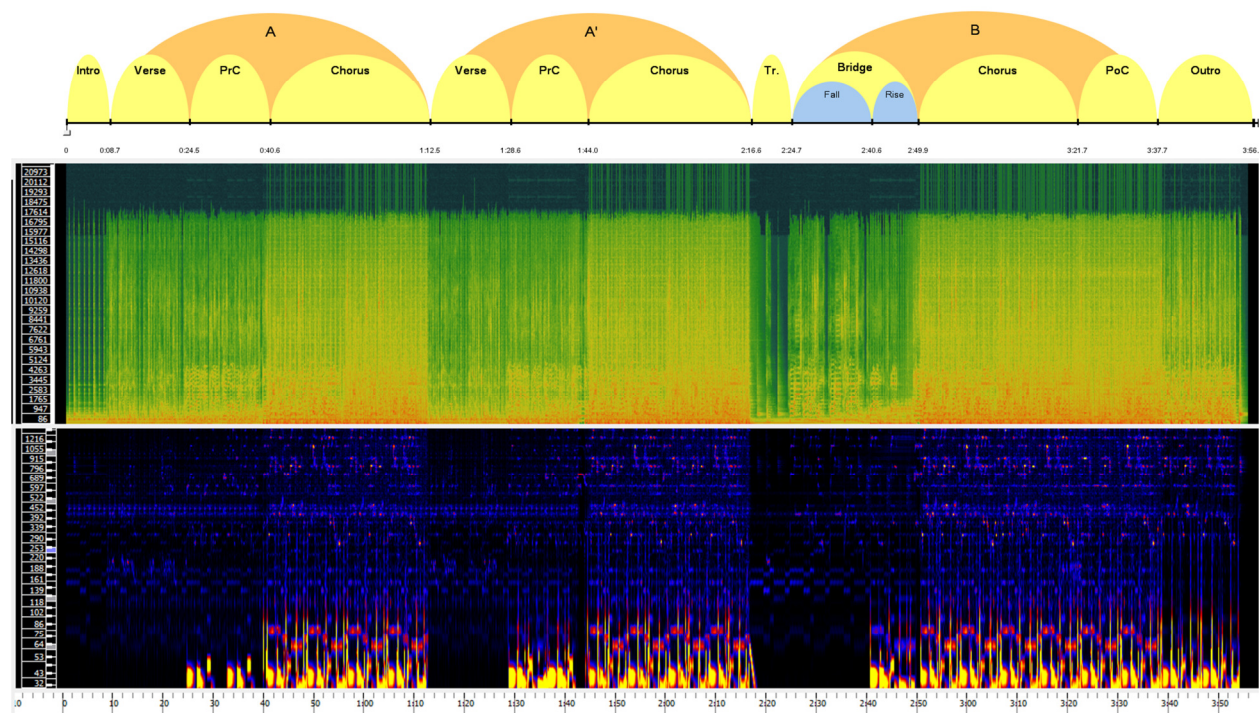


Example 5.2. *Full spectrograms and form timelines of recent pop songs.*

b) Katy Perry, “Roar” (2013).



c) Tove Lo, “Talking Body” (2014).



Example 5.2. *Cont.*

The above examples, of course, are chosen to make a point and do not represent all songs in either genre, but they do represent a substantial trend. In 20th-century popular music, changes in sonic density between sections are relatively minor compared to the much more explicit sonic changes in most recent pop songs. There are three reasons for this. First, songs in the 20th century were typically recorded using live instruments rather than computer-generated sounds and samples, and therefore the control over the spectral makeup of each sound was relatively limited. Second, the ability to simultaneously automate multiple timbral processes before the emergence of the DAW was much more limited than it is today. Third, different sections in a song were usually articulated by harmonic change, and therefore did not necessitate major changes in timbre.¹⁸ For example, in “Freedom 90,” the verse is heard over a V-IV-I chord loop, the pre-chorus over a C-minor vamp, and the chorus over a I-bVII-IV-I chord loop. Each section also features its own distinct melody. These harmonic and melodic shifts construct a clearly discernible structure. The timbre and texture, however, remain more-or-less the same in all three sections. The vocals, drums, piano, guitar, and bass are heard consistently throughout the entire song with no major balance changes. The only exceptions are the transition and bridge (4:22-5:25), in which the sonic density temporarily drops, as is visible in the spectrogram. The same pattern occurs in “Sweet Child O’ Mine” and “Crazy Little Thing Called Love.” The form is articulated via changes in harmony and melody, but the sonic density remains relatively stable throughout the song, except in transitional sections.

¹⁸ There may be a fourth reason for inter-section timbral changes being less pronounced in 20th-century popular music (especially in rock): As opposed to today’s pop, where producers, songwriters, and engineers design the song’s sonic trajectory and either create the sounds using a DAW or hire musicians to play predetermined parts, 20th-century rock bands were comprised of 4-5 members who often wrote and arranged the songs in collaborations. As a guitarist who played in rock bands in the 1990s, I can testify that at least in the situations that I have been a part of, every member of the band wanted to be heard in every section of a song. Ego was part of the equation, and musicians wanted to be heard as much as possible throughout the entire song.

By contrast, the chord loops in all three songs from Example 5.2 remain unchanged from beginning to end, save for the short instrumental intro to the chorus in “Cool for the Summer” (0:44-0:53 and 1:45-1:54). The form is defined in part by changes in the melody and the lyrics, but no less substantial are the changes in sonic density. For instance, the bass in “Talking Body” is initially withheld, entering only at the beginning of the pre-chorus (0:24). The presence of the sub-bass and its harmonic partials is significant, as reflected in Example 5.2.c. Even more substantial is the increase in sonic density in the chorus (0:40-1:13), as the drums finally enter and the sub-bass is boosted by an additional dirty bass sound.

The contrast between the spectrograms representing 20th-century popular songs and those representing current pop songs is in part a reflection of the dominance of the sonic dimension in shaping the form in today’s pop. The contrast between a verse and a chorus is usually characterized by a significant difference in sonic density. Transitional sections, such as the pre-chorus or the bridge, are partially defined by tension-increasing sonic alterations, as detailed in Chapter 4. The ability of today’s producers not only to choose which instruments/sounds will participate in the mix at any given moment, but also to determine precisely how much presence each sound will have in any range of the sonic spectrum, has given this music a new identity.

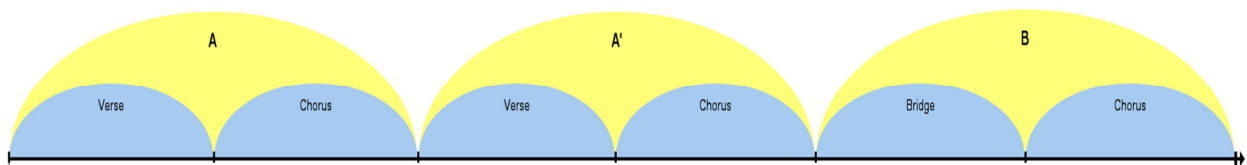
VERSE-CHORUS FORM

Verse-chorus form is by far the most commonly used song form in recorded popular music since the mid-20th century. Its two primary components are traditionally defined by their melodic content and lyrics. The different iterations of the verse share the same melody, but the lyrics change from one to the next. The chorus, on the other hand, is completely restated (melody

and lyrics) several times throughout the song. Because it is repeated multiple times, the main point, or “hook,” of the song will usually be found in the chorus. The verse’s role is to set up the arrival of the chorus by providing details/context in the lyrics that clarify the chorus’s hook. At minimum, a typical verse-chorus song will consist of three iterations of the verse-chorus pair, or two iterations of the pair followed by a bridge and a closing chorus.

The bridge evolved from the 32-bar **AABA** form. In fact, in the United Kingdom it is referred to primarily as a “middle-8,” due to its origin as the **B**-section of the 32-bar song form. In verse-chorus form, the bridge is a contrasting section that typically features new lyrics and melody and transitions into the final chorus. At times, it is replaced by other contrasting sections. For instance, in 20th-century rock songs, a guitar solo often takes the place of the bridge, and in some current pop songs, it can be omitted in favor of a rap verse by a guest artist.¹⁹

The verses and the bridge, each in its own way, set up the arrival of the chorus. Therefore, each of the verse-chorus and bridge-chorus pairs forms a section in a larger **AA'B** form, as shown in Example 5.3. This structure, as I mentioned, is the “bare bones” of verse-chorus form, excluding intros, outros, and other optional sections. Most songs feature a more elaborate form, but it is usually based on this structure.



Example 5.3. *Verse-Chorus form as AA'B.*

¹⁹ Each of these additions can either replace the bridge or precede it.

In the 1960s, a new section emerged – the pre-chorus (Summach 2011, par. 1). The pre-chorus was designed as a transition (usually harmonic) between the verse and the chorus. It contributed to the clarification of verse-chorus form's syntax – the verse was the *setup*, the pre-chorus was the *buildup*, and the chorus was the musical climax, or *peak*, roles that were defined primarily by harmonic function. Today, the pre-chorus is widely heard in pop songs, but the transition is not necessarily expressed by harmony. Rather, it is typically characterized by transitional sonic alterations such as adding or removing sonic layers.

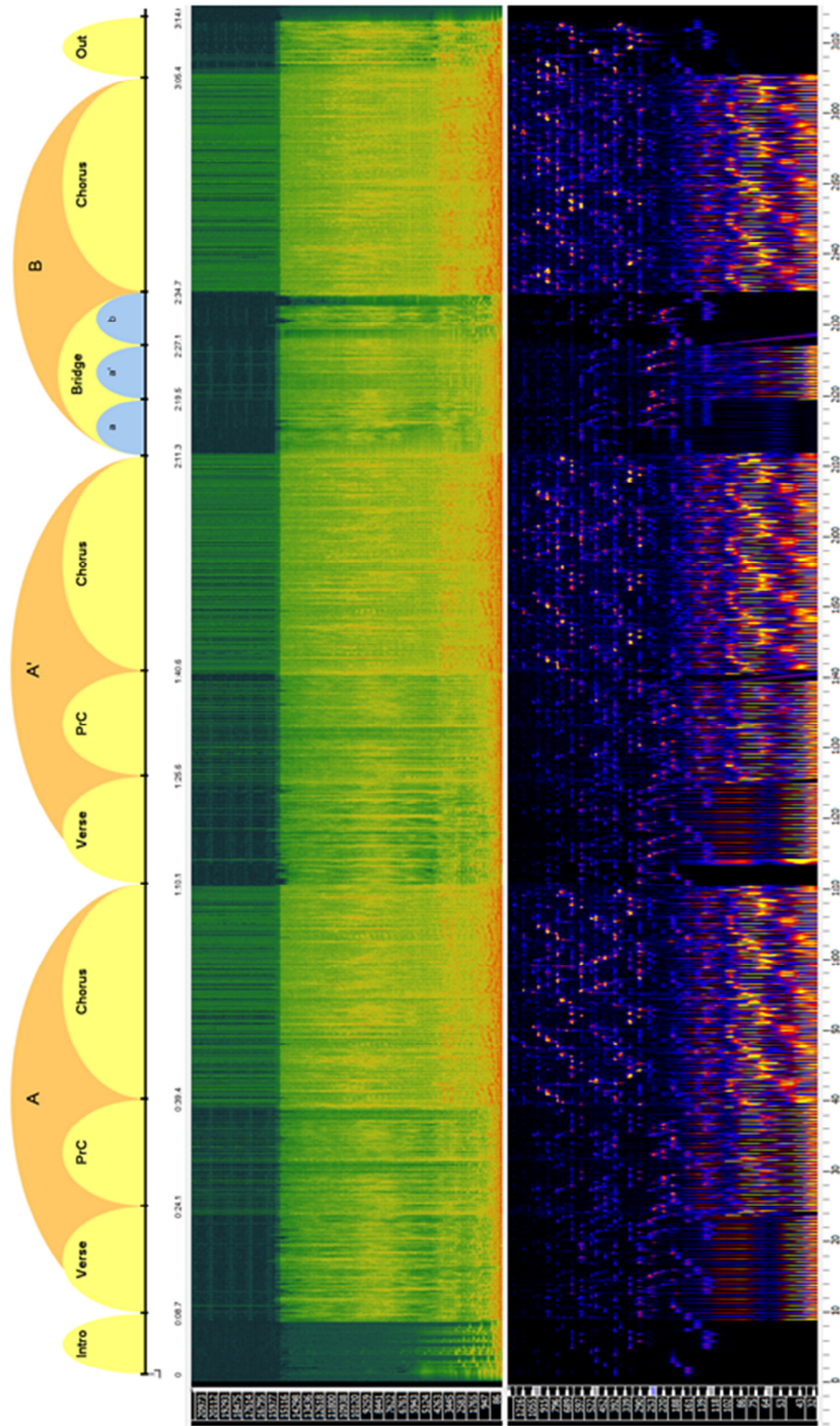
The sonic dimension in today's pop songs functions on a similar pattern. Although occasionally there are additional interim stages, the sonic arc of a typical pop song, on multiple structural levels, can be described as containing a setup, buildup, and peak. In most songs, these correlate with the traditionally defined verse, pre-chorus, and chorus, respectively. However, there is a growing practice in which the two dimensions diverge. In some songs, the buildup occurs in the first half of the chorus (following a prolonged setup in the verse and pre-chorus) and the peak arrives in the middle of the chorus. In others, the buildup persists throughout the entire chorus and the peak only arrives at the post-chorus. This trend is in part a result of the growing influence of EDM, where the buildup of sonic energy is prolonged and the arrival of a peak is often delayed.

The Setup-Buildup-Peak Sonic Sequence

Ariana Grande's "One Last Time" (2014) is an example of a song in which the traditional and sonic dimensions coincide. Example 5.4 shows a form timeline, detailing the sections as they would be traditionally defined, positioned above a plain and melodic range spectrogram. We can see that the first two large sections (**A** and **A'**) each feature a setup, a buildup, and a peak. In **A**, the setup in the verse (0:08-0:24) includes the vocal part, a soft accompanying synth riff, and a

four-on-the-floor (repeated quarter notes) kick drum. In the pre-chorus (0:24-0:39), the drum line intensifies to include mid- and low-toms, and both the vocals and the synth riff are doubled an octave above and enhanced with a reverb effect. This constitutes a buildup, as the sonic energy progressively increases toward the peak. The peak arrives at the chorus (0:39-1:10), as the synth riff expands to the bass, and a white-noise track (or perhaps it is an additional, “buzzy” synthesizer) saturates the sonic spectrum.

This sonic progression repeats itself in A’, but with a somewhat-higher degree of intensity from beginning to end. The second verse (1:10-1:25) is tenser than the first verse for two reasons. First, it begins by completely removing the drums, dropping the sonic energy below the baseline level set in the first verse, and generating anticipation for their return. Second, when the four-on-the-floor kick drum returns, it is doubled by a soft shaker, increasing its presence in the mid- and high-frequency range. In addition, the vocals are sparsely harmonized in the second half of this verse, and there is a short noise sweep preparing the arrival of the pre-chorus. Continuing on this path, the sonic spectrum in the second pre-chorus (1:25-1:40) is denser than the first, with cymbals and claps added to the texture, as well as subtle background vocals that gradually become more audible, creating a pseudo-sweep in anticipation of the chorus. Since this buildup is more intense than the previous one, the level of energy in the second chorus (1:40-2:11) is predictably higher than it was in the first. The vocals are harmonized throughout and the white noise track increases its presence in the mix via amplification.



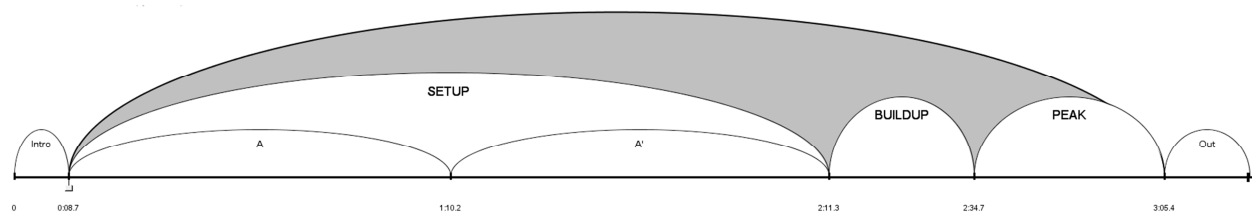
Example 5.4. Ariana Grande, "One Last Time" (2014).

The **B**-section seemingly has only a buildup and peak, with no setup. The bridge, by its nature, is a transitional section, usually acting as a more-intense version of the pre-chorus. However, it has an internal structure which includes a setup and a buildup that prepare the arrival of the peak/chorus. The first part of the bridge, labeled as **a** (2:11-2:19), increasing tension due to the significant drop in sonic energy, acts as the internal setup. It is followed by a buildup in **a'** (2:19-2:27), formed by the intensifying drums, a rhythmic noise pulse, an octave doubling of the synth melody, and especially a noise sweep toward the end of the subsection. This, however, turns out to be a “deceptive” buildup, as instead of arriving at the peak, the sonic energy once again abruptly drops in **b** (2:27-2:35), leaving only the vocal melody and the contrapuntal synth line. This drop in energy further increases the tension, as we anticipate the delayed arrival of the peak. A short noise sweep at the end of **b** leads into the final chorus, which is the most climactic of the three, as is idiomatic in this genre. The lead vocal melody in this chorus is not only harmonized, but is also supplemented with an additional contrapuntal voice in the upper register, and the texture (particularly the noise fill) becomes brighter overall as the high frequency range is given an EQ boost.

Zooming Out and Alternative Interpretations

I mentioned above that the bridge is a transitional section. Internally, it has a setup and a buildup, but if we consider what came before it, the setup is highly anticipatory due to the abrupt gap in energy created by the sonic fall. Taken in the context of the entire song, therefore, the bridge constitutes the most intense buildup of tension in this song (as well as in most songs in this genre). The large sections that came before it, **A** and **A'**, cover a lot of ground in terms of accumulating sonic energy. However, this accumulation occurs over a prolonged period of time, and therefore these large sections are more stable relative to the bridge. Moreover, since each

new peak is higher in sonic energy than the one that preceded it, the large-scale “goal” in this song, and others, is the third chorus. Therefore, on the deepest structural level, **A** and **A'** can be interpreted as a combined large-scale setup, while the bridge and the final chorus are respectively the large scale buildup and peak, as illustrated in Example 5.5.



Example 5.5. *Ariana Grande, “One Last Time.” Sonic reduction.*

This sonic reduction challenges my own initial interpretation of this form as consisting of three distinct large sections (**A**, **A'**, and **B**). On the one hand, I lump two of these sections into a single, large-scale setup, and on the other hand, I split the third large section into two. On its face, this seems like a contradiction. However, my listening experience points to the simultaneous existence of these two interpretations. On a local level, I hear the bridge and the third chorus as a pair that features internal interaction. On the deeper structural level, I hear the bridge as transitioning from the previous material to the final peak, similar to a pre-chorus transitioning from a verse to a chorus on a local level. While the former is a more dominant mode of hearing than the latter, following the deeper structure adds an interesting dimension to understanding the expression of verse-chorus form in this genre.

Intro and Outro

So far in my analysis of “One Last Time,” I have ignored the intro, because it is outside the core sonic trajectory in verse-chorus form. This does not diminish its significance. Almost every pop song starts with an intro that sets up the arrival of the first verse, introduces thematic material, and allows for a smooth opening. Just as importantly, the arrival at the verse sets up the baseline energy of the song, which affects the level of tension produced when the sonic energy intensifies or drops.

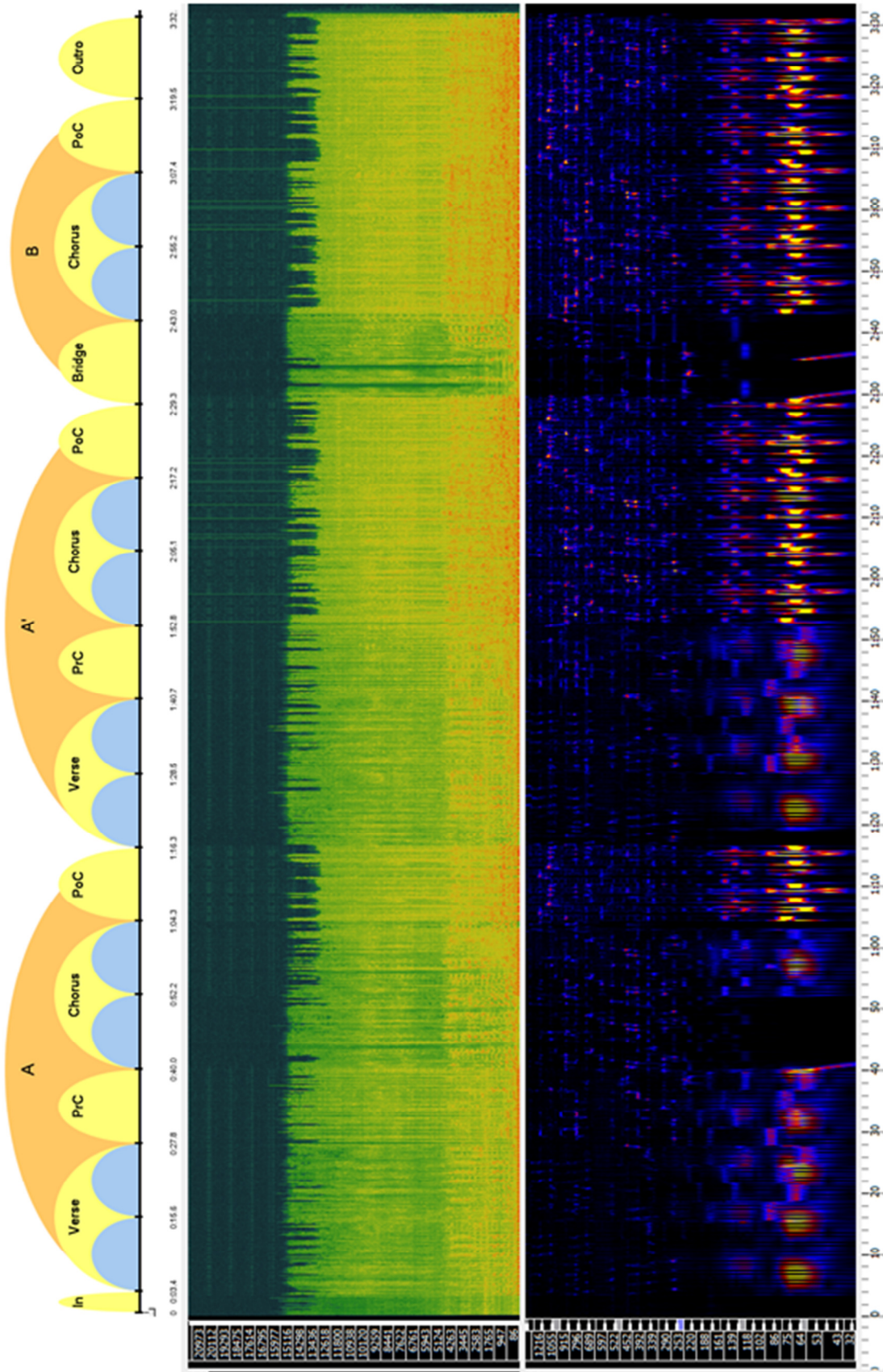
The outro, too, is outside the core trajectory. In “One Last Time,” it functions as a pseudo-coda in the sense that it recapitulates the melodic and lyrical material from the chorus in a mellower manner. However, unlike a real coda, it does not reinforce closure, since this song, like most songs in this genre, does not achieve complete closure. Outros can assume a number of forms. For example, in Jack Ü’s “Where Are Ü Now” (2015), the outro simply consists of delayed echoes from the final chorus. In Robin Thicke’s “Back Together” (2015) the music stops abruptly and Thicke’s voice trails out to end the song. Selena Gomez’s “Same Old Love” (2015) ends with a single piano chord. Whether or not there is an actual ending section, such as in “One Last Time,” or simply a trailing sound, songs in this genre almost always conclude on an open-ended gesture or note, as detailed in Chapters 3 and 4.

Diverging Dimensions

While the traditional and sonic dimensions are in sync through the entirety of “One Last Time,” in more and more songs the setup-buildup-peak scheme diverges from the traditional trajectory. The most common divergence of the sonic dimension from the traditional one is a buildup that begins in the chorus rather than in the pre-chorus. The buildup, in these cases, can

arrive at the peak midway through the chorus, or persist throughout the entire chorus, in which case there will be a post-chorus carrying the role of the peak section. In some songs the peak is completely avoided, but this is quite rare.

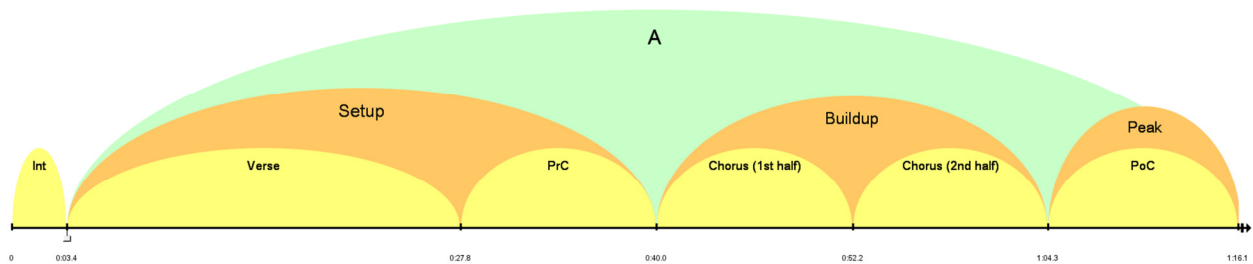
In-chorus buildups normally begin with a dip in sonic energy in order to heighten the tension and contrast with the preceding pre-chorus. One such example can be heard in Taylor Swift's "I Knew You Were Trouble" (2012). Until the arrival of the first chorus, the verse and the pre-chorus consist of a stepwise sonic buildup (0:03-0:40). A noise sweep leads from the pre-chorus to the chorus, leading the listener to expect a climax, but as shown in Example 5.6, there is instead a sharp fall in sonic energy at 0:40. The drums and bass drop out, and only a guitar accompanies Swift's voice. The tension intensifies again in the second half of the chorus (0:52) – the drums and bass return, but they are initially subdued by an EQ, and are gradually released by a filter sweep, while simultaneously increasing their rhythmic intensity.



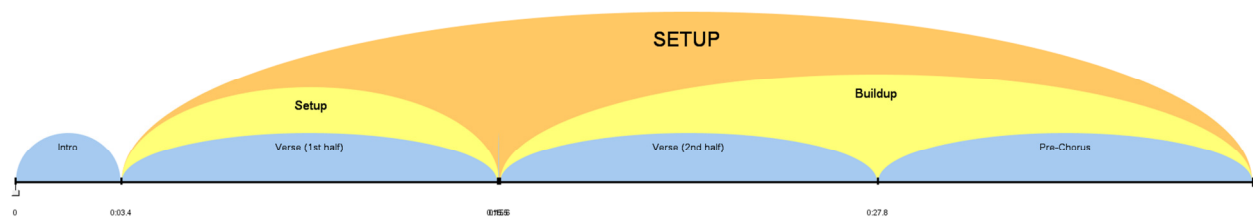
Example 5.6. Taylor Swift, "I Knew You Were Trouble" (2012).

This in-chorus buildup redefines the role of the pre-chorus in the first major sonic cycle in this song. When reducing the four sections of **A** (verse, pre-chorus, chorus, and post-chorus) to reflect the setup-buildup-peak paradigm, the verse and pre-chorus constitute the setup section, as shown in Example 5.7.a. This is because the in-chorus buildup is significantly more intense than the preceding sequence, which is more stable. However, when zooming in to the local level, the internal buildup within the setup section becomes apparent, as shown in Example 5.7.b. The “real” setup occurs only in the first half of the verse. The second half of the verse and the pre-chorus incrementally build up sonic energy that aims to climax at the chorus, but is ultimately interrupted by the abrupt fall in the chorus.

a) Setup-buildup-peak sequence in A, 0:03-1:16.

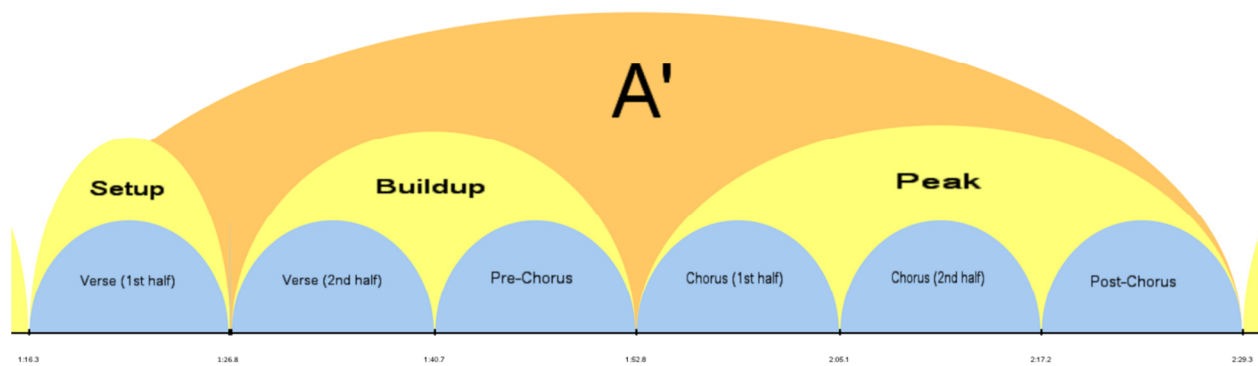


b) Internal development within the larger setup section, 0:03-0:40.



Example 5.7. Taylor Swift, “I Knew You Were Trouble.”

This internal buildup in the verse and pre-chorus is particularly worth noting, since in the next large section, **A'**, it indeed fulfills its aim and peaks at the chorus, rather than being interrupted by an additional buildup. This causes the sonic sequence of **A'** to shift – the first half of the verse (1:16-1:28) is now the setup, the second half of the verse (1:28-1:40) and the pre-chorus (1:40-1:52) form the buildup, and the chorus (1:52-2:17) and post-chorus (2:17-2:29) constitute the peak, as shown in example 5.8.



Example 5.8. Taylor Swift, “I Knew You Were Trouble,” 1:16-2:29.

In the final large section **B**, the chorus (2:43-3:07) again marks the arrival of the peak. This is unsurprising for two reasons. First, since the interruption that occurs in **A** is not repeated in **A'**, there is no clear expectation that it will be repeated in **B**. Second, it would be more difficult to have an in-chorus buildup when the chorus follows a bridge, since the bridge already represents the most intense buildup in most songs. In this song specifically, while the bridge does not feature a section-long filter sweep or other ascending gesture, the dip in sonic energy is so stark that the tension is significantly heightened, preparing the listener for the most climactic

peak in the song. Indeed, most songs in which there is an in-chorus buildup in **A** and/or **A'**, do not have one in **B**, and use the final chorus as a sonic peak. Additional examples of this can be heard in Justin Bieber's "Beauty and a Beat" (2012) and Britney Spears's "Hold it Against Me" (2011).

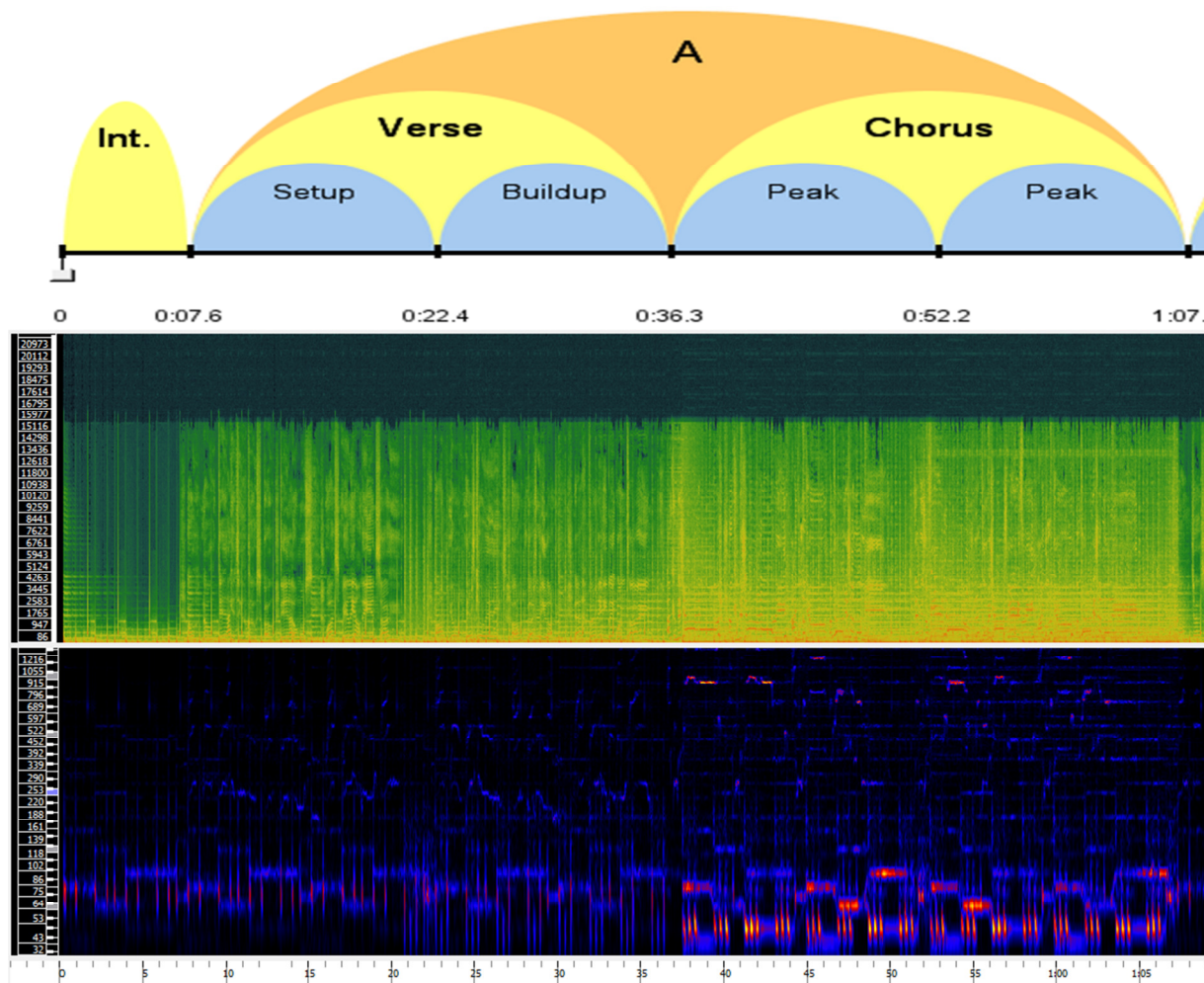
As in "One Last Time," the third and final peak is the most sonically-dense one. The gap in sonic density between the second chorus and the bridge is the largest between any two sections until that point, and the sonic ground covered between the bridge and the third chorus is even larger (the largest in the entire song). This makes the bridge a pivotal point in each of these songs (as well as most verse-chorus songs in this genre), not only in the traditional dimension, but also (and perhaps primarily) in the sonic dimension.²⁰ The sonic "fall and rise," on the other hand, occurs in the bridge of almost every verse-chorus song in this genre. For example, the fall/rise sequence is present without changing the chord loop in the aforementioned "Talking Body" (2:24-2:50) and "Cool for the Summer" (2:19-2:49), as well as in numerous other songs, including Calvin Harris's "Outside" (2014, 2:39-2:55), Sia's "Elastic Heart" (2013, 2:54-3:08), and Carly Rae Jepsen's "Call Me Maybe" (2012, 2:15-2:32).

In-Verse Buildups

As was the case in "I Knew You Were Trouble," sonic energy can begin to build up during the verse, when sonic layers are added in the second half, beginning a stepwise buildup. This buildup normally persists through the pre-chorus, but in songs where there is no pre-chorus, the entire buildup may be contained in the verse. Such an in-verse buildup can be heard in the **A** section of Katy Perry's "Unconditionally" (2013). The sonic process centers on expanding the

²⁰ Although both of these songs feature a change in the chord loop in the bridge, such a change is not a defining characteristic of the bridge in contemporary pop music.

frequency presence from the low to the high range. The setup in the first half of the verse (0:07-0:22) is characterized by a generally muffled sound; only Perry's voice, a woodblock, and sparse whistle-like sounds break out of the low frequency range. In the second half of the verse (0:22-0:36), there is an immediate increase of presence in the mid-frequency area, due to added guitars and drums, which indicates a stepwise sonic climb. This subsection is further intensified by an EQ gradually amplifying the higher frequencies, as shown in Example 5.9 (particularly between ~1,000 and ~5000 Hz). To complete the buildup, there is a partial break leading up to the chorus, in which a crescendo in the cymbals is the only accompaniment under the vocal pickup to the chorus.



Example 5.9. Katy Perry, “Unconditionally” (2013), 0:00-1:07.

The buildup in the setup-buildup-peak sequence is a syntactic function, and does not necessarily have to be represented by an entire section or subsection. For instance, in Tove Lo’s “Habits” (2014) not only is there no pre-chorus, but there is also no in-verse buildup comparable to the one in “Unconditionally.” The first verse (0:09-0:44) acts as a setup with no sonic development throughout, and is immediately followed by the chorus (0:44-1:19). However, in the last measure of the verse (0:42-0:44), there is a break and a very high pitch bend accompanying Lo’s pickup. This brief gesture heightens the sense of anticipation and is enough

to fulfill the function of the buildup even without a section- or subsection-long “warning” that a climax is approaching. This is not very common, as most recent pop songs do have a pre-chorus, and those that do not have one typically feature some sort of sonic development within the verse. However, a fulfillment of the buildup function by a short gesture without a pre-chorus is completely congruent with the syntactical logic of this genre, and does not sound unusual.

DISTINGUISHING BETWEEN THE POST-CHORUS, BRIDGE, AND TRANSITION

Earlier, I discussed the essential sections of verse-chorus form in today’s pop music, including the bridge and the almost essential pre-chorus. In my analyses, I have also mentioned sections that elaborate this basic form, such as transitions and post-choruses. Although, in the “bare bones” version of verse-chorus form, the only section that separates the second and third choruses is the bridge, in practice there are a number of types of sections that can be added in this temporal space. In order to better understand the sonic behavior of sections that elaborate the basic verse-chorus form, it is important to differentiate between them.

The second chorus (i.e., the chorus in **A'**) can be followed by a bridge, a transition, a post-chorus, a third verse, a rap verse (in songs that feature a guest rapper), or some kind of dance-oriented breakdown. I will touch only briefly on the latter three, as their properties are mostly self-explanatory. However, informal conversations I have had with a number of popular music scholars indicate that the distinctions between the post-chorus, transition, and bridge in this music is not clear, particularly in the absence of harmonic changes. I will therefore begin by defining the characteristics of each of these sections in the context of contemporary pop music.

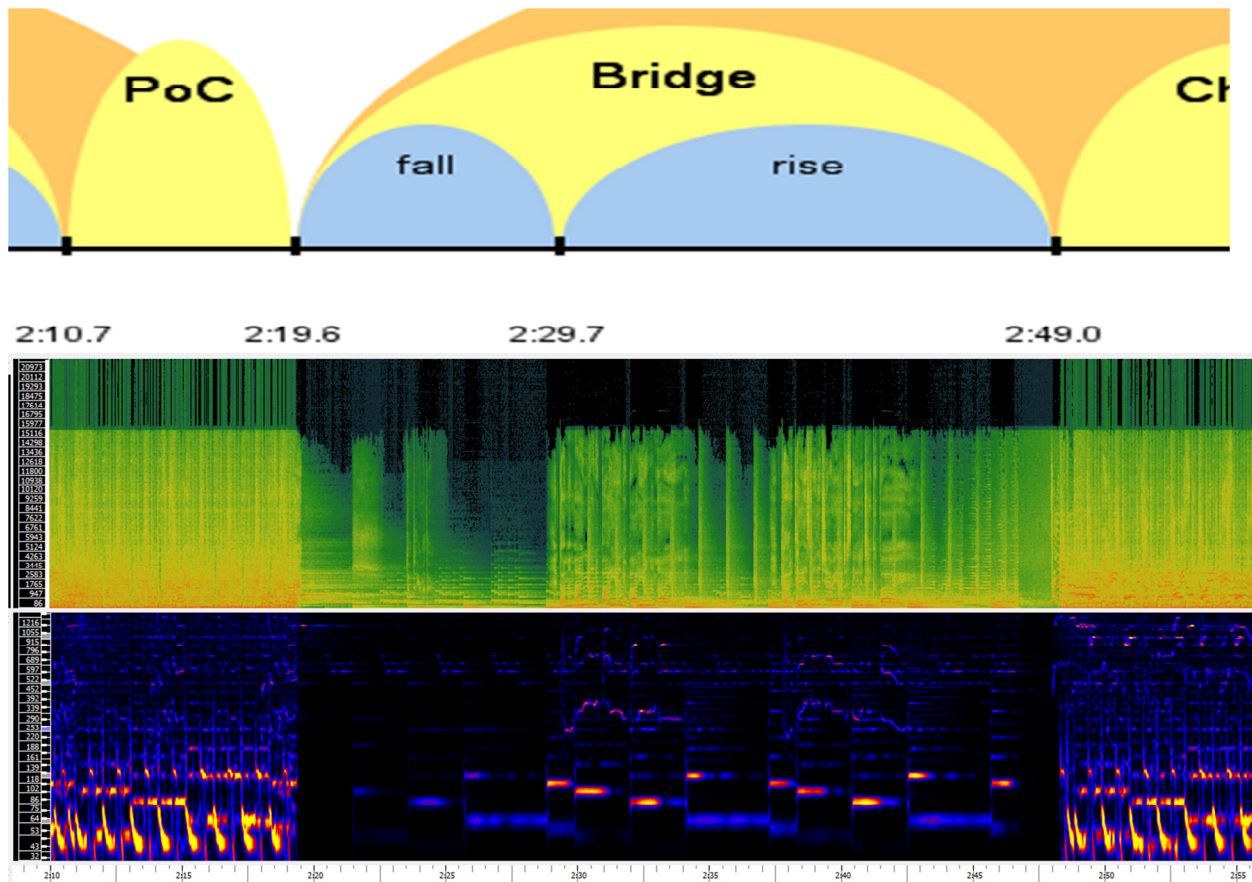
Bridge

The bridge is most commonly comprised of two parts,²¹ which I term the *fall* and the *rise*, and it features some form of development or fragmentation of previous melodic and lyrical material, similar to the traditional development section. As their names suggest, the fall is characterized by a sharp drop in sonic energy, and the rise by what is usually the most intense buildup in a song, leading up to the final peak/chorus. The rise can take the form of a *riser* in EDM-oriented songs, but the two are not synonymous. A riser, as previously described, is a very intense dance buildup that normally includes one or more of the following: ascending sweeping gestures, drum intensification, and a break (at its end). The term “rise” refers to a subsection of the bridge (usually its second half), in which the music intensifies significantly, but this intensification can be achieved via gestures that are not necessarily risers.

All three of the initial examples of recent pop songs from the beginning of this chapter include a bridge comprised of a fall and a rise. Example 5.10 shows a spectrogram of the bridge (2:19-2:49) in “Cool for the Summer.” Following a climactic post-chorus, the sonic energy drops to include only a soft piano riff and a gentle bass sound accompanying Demi Lovato, as she whispers a fragment of earlier lyrics (“Don’t tell your mother”). This low-energy subsection constitutes the fall (2:19-2:30). In the rise (2:30-2:49), Lovato returns to her normal singing voice, repeating the lyrics and melody from the end of the verse in a mellower tone, and the sonic energy steadily increases: the bass gradually becomes more aggressive and “dirty”; an additional quickly-arpeggiating piano is added in the background; reverb is progressively added to give the general sound more presence; and, starting at 2:42, the piano riff is mirrored by digitally-reversed plucked strings, which gradually move to the foreground via a crescendo. The

²¹ Two previous examples featured bridges that were not in two parts. “One Last Time” had a three-part bridge and “I Knew You Were Trouble” had an indivisible bridge. However, these do not represent the norm.

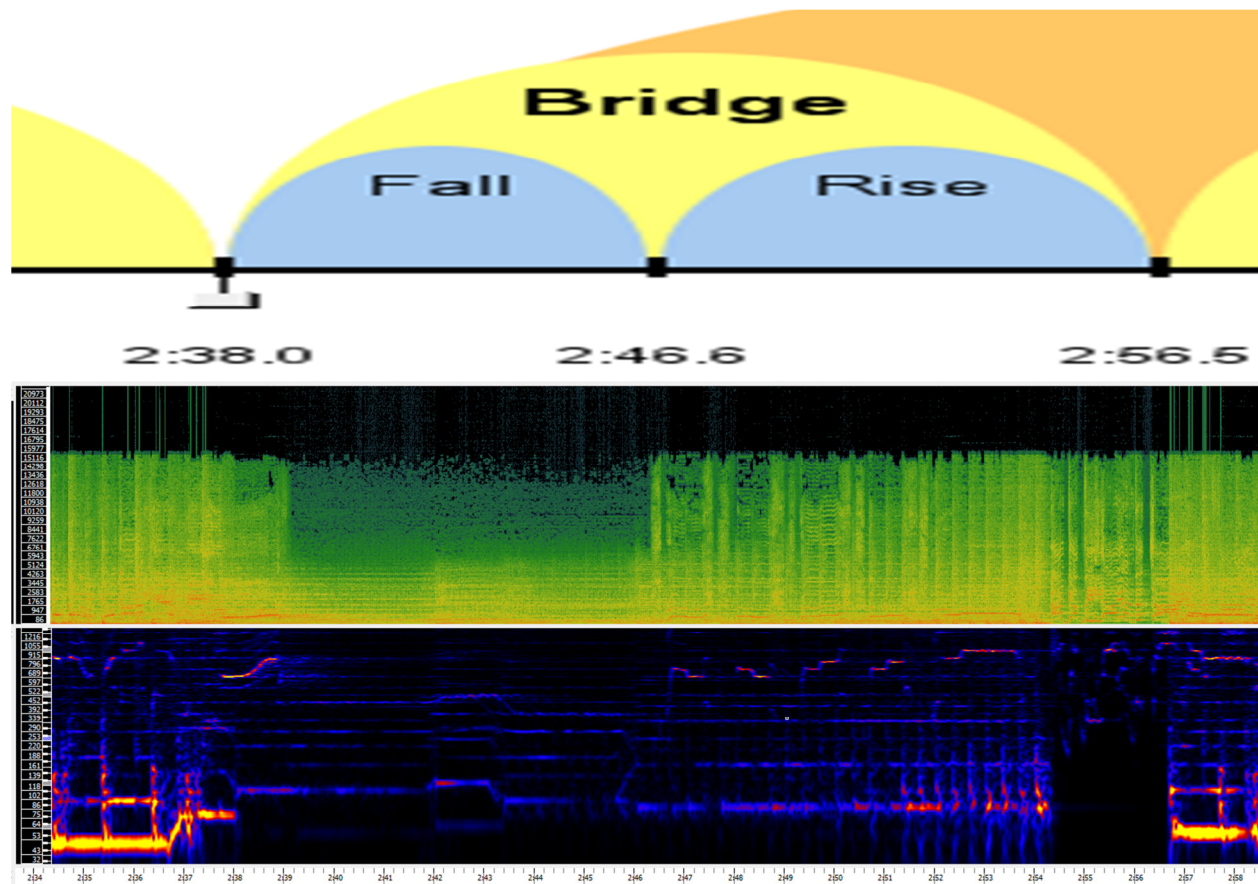
rise is interrupted by a break (2:47) that propels the music into the final chorus/climax as white noise sweeps into the sonic space.



Example 5.10. Demi Lovato, “Cool for the Summer,” 2:10-2:55.

In “Roar” (Example 5.11), the fall-rise process is simpler (2:38-2:56). A synth pad that was previously in the background is suddenly exposed in the foreground, as all other instruments drop out (2:38). At 2:44, the kick drum, snare drum, and bass almost stealthily return, gradually growing louder in preparation for the rise, which begins at 2:46. The rhythm section (drums and

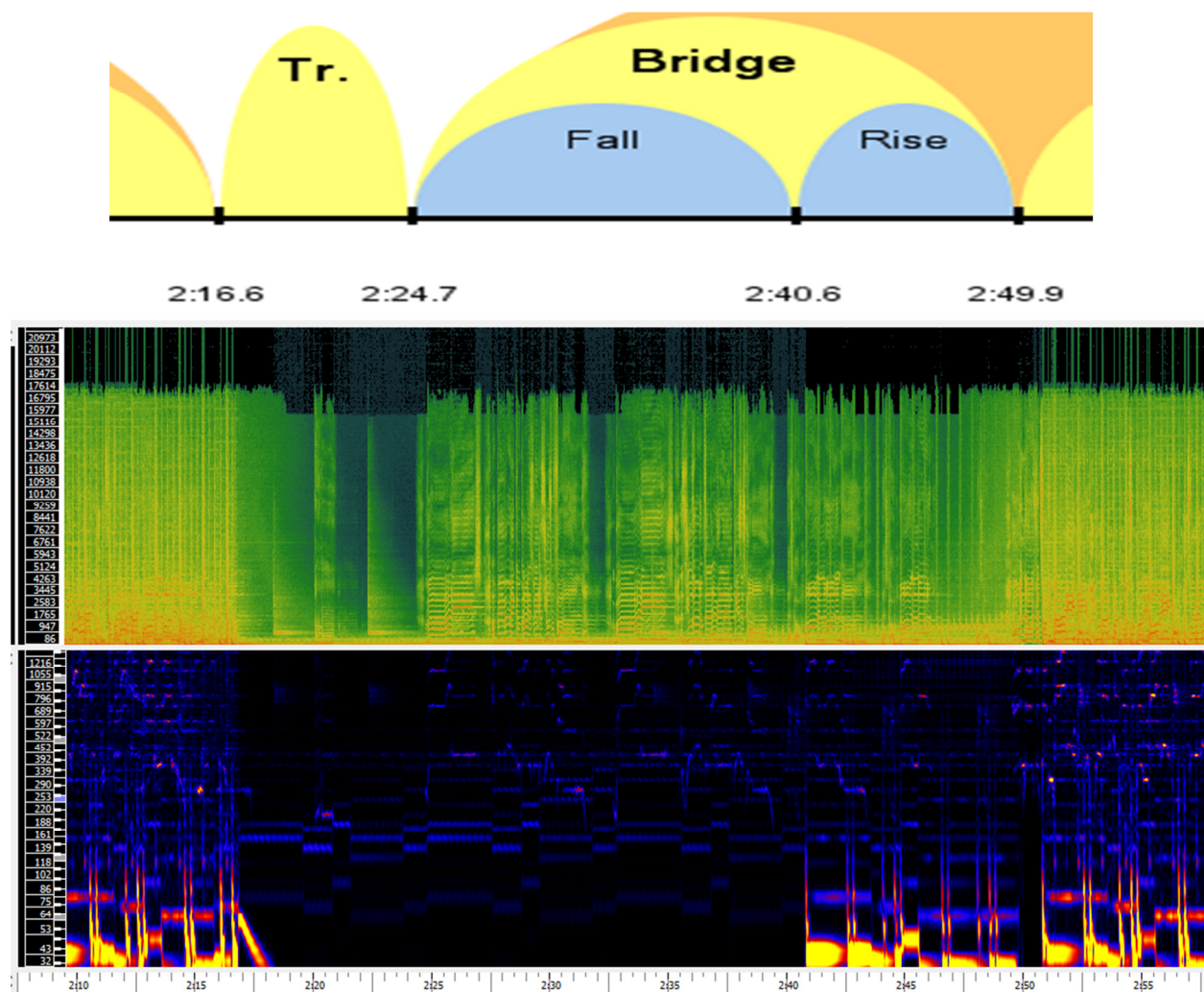
bass) continue to crescendo, rapidly moving into the foreground. Fragmentation of the melody and lyrics begins as Perry repeatedly sings the title word “roar,” the last repetition ascending melodically toward the return of the chorus. As in the previous example, this rise is sealed with a break, followed by the arrival of the final chorus.



Example 5.11. *Katy Perry, “Roar,” 2:34-2:58.*

In “Talking Body,” the drop in sonic energy occurs in a transition (2:16-2:24). This transition is not yet a fall, since there is no development or fragmentation of previous melodic or lyrical material. Rather, it serves as an intro to the bridge (2:24-2:49), as shown in Example 5.12.

The fall (2:24-2:40) introduces new lyrics and melody based on the plural form of the title word “body.” Toward its end, a percussive, rhythmic sound is added in anticipation of the rise (similarly to the rhythm section in “Roar”). Fragmentation of the chorus lyrics (“On and on and on”) takes place in the rise (2:40-2:49), as the percussive sound becomes progressively louder, and a noise sweep anticipates the final chorus.



Example 5.12. Tove Lo, “Talking Body,” 2:10-2:55.

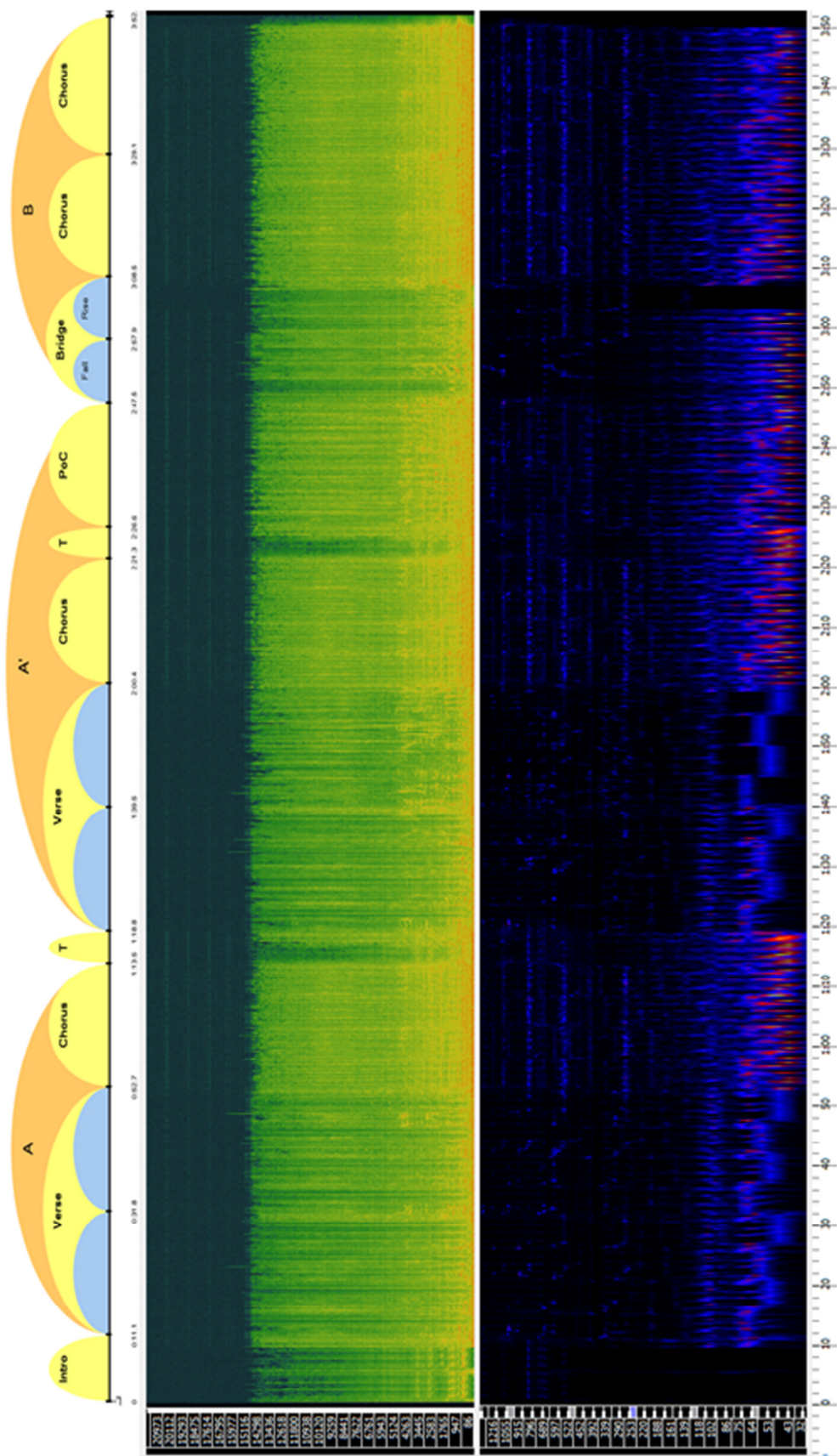
As these examples indicate, the bridge constitutes the most significant structural buildup in the sonic dimension of verse-chorus form. Therefore, it is almost always the case that the fall represents the lowest level of sonic energy in the song and the following peak represents the highest, maximizing the range of the energetic buildup.

Transition

A transition may initially sound like the fall in the bridge, due to its sonic drop and detachment from the peak (chorus or post-chorus) that preceded it, as was the case in “Talking Body.” However, it features two key differences from the fall. First, the transition contains no development or fragmentation of previous material. It is usually a simple exposed accompaniment with no vocals, such as in Becky G’s “Shower” (2014, 2:03-2:11), a recurring hook, such as in Ariana Grande’s “Problem” (2014, 0:58-1:03 and 1:49-1:54) and Miley Cyrus’s “We Can’t Stop” (2013, 1:13-1:25 and 2:25-2:37), or simply a short interjection, such as in Ariana Grande’s “Focus” (2015, 2:06-2:09). Second, it is usually shorter than the typical section or subsection, and is comparable to an intro in duration. I consider transitions to be functionally equivalent to intros and outros – the intro transitions from silence to the first verse, and the outro transitions into silence.

A transition usually connects **A** and **A’** or **A’** and **B**, and is not an integral part of the setup-buildup-peak sequence. Depending on the nature of its sonic treatment, it is either a mini-setup or a mini-buildup to the “real” setup in the large section that follows it. The examples cited in the previous paragraph are all mini-setups, but a transition acting as a mini-buildup can be heard, for instance, in Katy Perry’s “I Kissed a Girl” (2008, 1:03-1:06), in which a crescendo of background vocal “aahs” lead to the second verse.

Of all the songs in my sample, I have only found one transition that is internal to a large section, in Taylor Swift's "Out of the Woods" (2014). As shown in Example 5.13, this transition (2:21-2:26) is located between the chorus and the post-chorus in **A'**. It is an exact repetition of a previous, external transition (1:13-1:18), which connects **A** and **A'**, and is a shorter and more intense version of the intro (0:00-0:11). This is an interesting transition – the plain spectrogram indicates a substantial drop in sonic density while the melodic range spectrogram shows increased intensity in the low frequency range (both in loudness and rhythmic rate), due to a frantic foregrounded kick drum. Perhaps because of its unusual intensity for a transition, the producers of this song (Swift, Max Martin, and Jack Antonoff), decided that it would be more appropriate as a setup to the post-chorus rather than to the bridge.



Example 5.13. Taylor Swift, "Out of the Woods" (2014).

Post-Chorus

The primary defining trait of the post-chorus is that it maintains or intensifies the sonic energy of the preceding chorus. It is either the point of arrival of the peak or its continuation. I hear most post-choruses as belonging to one of two major categories: the “regular” post-chorus and what I call the *dance post-chorus*.

The regular post-chorus is an additional section that follows the chorus, featuring a new melody, and as mentioned, maintaining or intensifying the sonic energy. It can be further divided into two subcategories: *separate* and *attached*. The separate post-chorus is an independent section, sonically and melodically distinct from the chorus. Examples of such post-choruses occur in Carly Rae Jepsen’s “Call Me Maybe” (1:59-2:16), Sia’s “Chandelier” (2014, 1:27-1:50 and 2:55-3:19), and Taylor Swift’s “You Belong with Me” (2009, 2:39-2:56). The attached post-chorus can be heard as an extension of the chorus rather than as its own section. For instance, in Selena Gomez’s “Good for You” (2015, 1:59-2:10), the sonic density increases slightly in the post-chorus, but not significantly enough to completely separate it from the chorus, and the vocal melody simply repeats the final gesture from the chorus. Other examples include Jason Derulo’s “Want to Want Me” (2015, 1:09-1:18, 2:08-2:16 and 3:07-3:23) and Ellie Goulding’s “Lights” (2010, 2:15-2:32).

The dance post-chorus is an EDM-influenced climactic section that follows an in-chorus buildup. In this type of post-chorus, the vocals are replaced by a high-energy synth riff or a dubstep-style breakdown. The preceding chorus will almost always build up sonic energy with an EDM-style riser, or at least some elements of it. I mentioned an example of such a post-chorus in Chapter 4 – the riser-drop sequence in Alesso’s “Heroes.” The buildup begins in the pre-chorus (0:22-0:38), significantly intensifies in the first half of the chorus (0:38-0:53), and

turns into a full-fledged riser in the second half of the chorus (0:53-1:08). The post-chorus (1:08-1:24) features a high-energy synth riff that is already present in the riser, but is initially in the background and gradually moves to the foreground with the aid of a filter sweep. Another very similar example can be heard in Calvin Harris's "Outside" (2014), where the buildup occurs throughout the chorus (0:39-1:09), and a previously present synth riff becomes the center of the climactic post-chorus (1:09-1:24).

Third Verse, Dance Breakdown, and Rap Verse

As I mentioned earlier, the properties of the third verse, dance breakdown, and rap verse are self-explanatory. However, I will nevertheless mention some examples of these sections.

A third verse is extremely rare in today's pop. There is only one song in my sample that includes one – Ludacris's "Representin'" (2012). It (2:36-3:04) replaces the bridge, eliminating the "twist in the plot." However, each large section in this song contains a post-chorus (1:14-1:25, 2:22-2:36, and 3:32-4:01), which partially compensates for the loss of variety.

The dance breakdown as a standalone section is also quite rare, since most climactic dance-oriented sections are dance post-choruses, as described above. However, there is at least one example of a dance breakdown in my sample that does not act as a post-chorus, but is instead an independent entity outside of the large sections. This is a dubstep-oriented breakdown that occurs in Britney Spears's "Hold it Against Me" (2011, 2:17-2:46). I do not interpret it as a post-chorus for two reasons. First, it is separated from the chorus by a transition (2:10-2:17); and second, its sonic character contrasts with the chorus. Though it is not impossible for an internal transition to be interjected between a chorus and a post-chorus, as was the case in "Out of the Woods," the sonic relationship between the chorus and the breakdown in "Hold it Against Me" is

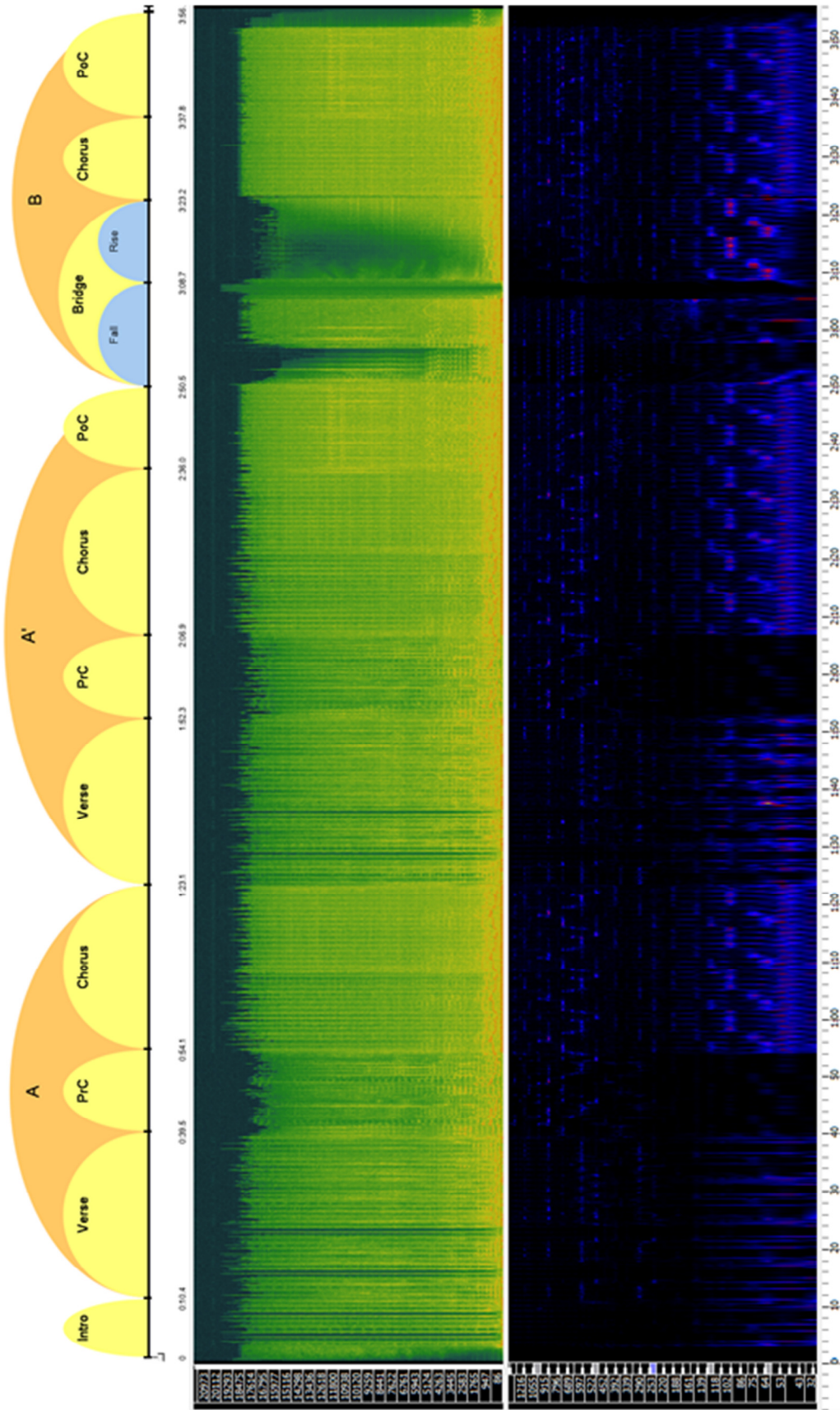
very different from that in “Out of the Woods.” In the latter, the sonic texture in both sections is nearly identical, and therefore the post-chorus sounds related to the chorus to a very high degree. In the former, however, the dance breakdown features a very aggressive dubstep bass, in stark contrast to the bright and clear texture of the chorus, rendering an interpretation of it as related to the chorus difficult to make.

As opposed to the third verse and the dance breakdown, the rap verse is quite common in today’s mainstream pop songs that feature a guest rapper. In that case, the rap verse will typically occur between the second and third chorus, either in place of the bridge or as a standalone section preceding the bridge. Examples of rap verses that replace the bridge can be heard in Justin Bieber’s “Beauty and a Beat” (featuring Nicki Minaj, 2:30-3:00), Katy Perry’s “California Gurls” (2010, featuring Snoop Dog, 2:26-3:05), and Fifth Harmony’s “Work from Home” (2016, featuring Ty Dolla \$ign, 2:19-2:56). By contrast, in Ariana Grande’s “Problem,” the rap verse (1:54-2:22), featuring Iggy Azalea, is a sonically detached section, interjected between the second chorus and the bridge. Similarly, Kendrick Lamar’s rap verse in Miguel’s “How Many Drinks” (2012, 2:11-2:57) is also succeeded by a bridge.

Considering the Sonic Dimension – Examples in a Wider Context

Till the World Ends

At this point, I would like to return to a song that I have mentioned a number of times: Britney Spears's "Till the World Ends" (2011). This is the first song I analyzed when I started working on this research project, and it particularly exemplifies the importance of the sonic dimension in songs in this genre. I have discussed this song on several occasions, both in the context of teaching and in conversations with other music theorists. The initial instinct of most of my conversation partners was to consider the section that immediately follows the second chorus (2:36-2:50) a bridge (see Example 5.14). Admittedly, this was my initial instinct as well. This section introduces new lyrics and melodic material following two verse—pre-chorus—chorus cycles. By all traditional definitions (with the exception of harmony – the chord loop, i-III-VI remains unchanged throughout the song), this is a bridge. However, when considering the sonic dimension and the genre, this is unquestionably a post-chorus. First, it maintains the level of sonic energy of the chorus and even slightly intensifies it by boosting the high frequency range. Second, the section that follows it is undoubtedly a bridge. It features a fall (2:50-3:08), in which the sonic energy drops and there is fragmentation of previous melodic and lyrical materials, followed by a rise (3:08-3:23), in the form of the full-mix filter sweep discussed earlier.



Example 5.14. Brimey Spears, “Till the World Ends” (2011).

Another thing to notice in this song is the recycling of melodic materials by changing the sonic context. For instance, the melody first heard in the pre-chorus (0:39-0:54) is repeated in the chorus (0:54-1:23). While the lyrics change, it is the sonic texture that defines these sections as distinct (imagine an acoustic version of this song – the pre-chorus and the chorus would likely be perceived as a unified section). Moreover, the chorus is repeated – lyrics, melody, and accompaniment – in the rise of the bridge, the only modifier being the filter sweep, transforming its function from a sonic peak to a buildup. In total, out of 3 minutes and 56 seconds, the pre-chorus/chorus melody is in the foreground for 1 minute and 57 second, almost exactly half the duration of this song (or exactly half, if we exclude the 2-second metallic sweep leading up to the intro). Furthermore, the chorus vocals are juxtaposed in the background during the final pre-chorus (3:37-3:56), bringing the total presence of this melody to 2 minutes and 16 seconds. This is a striking example of sonic treatment enabling thematic development without changing either the melody or the harmony.

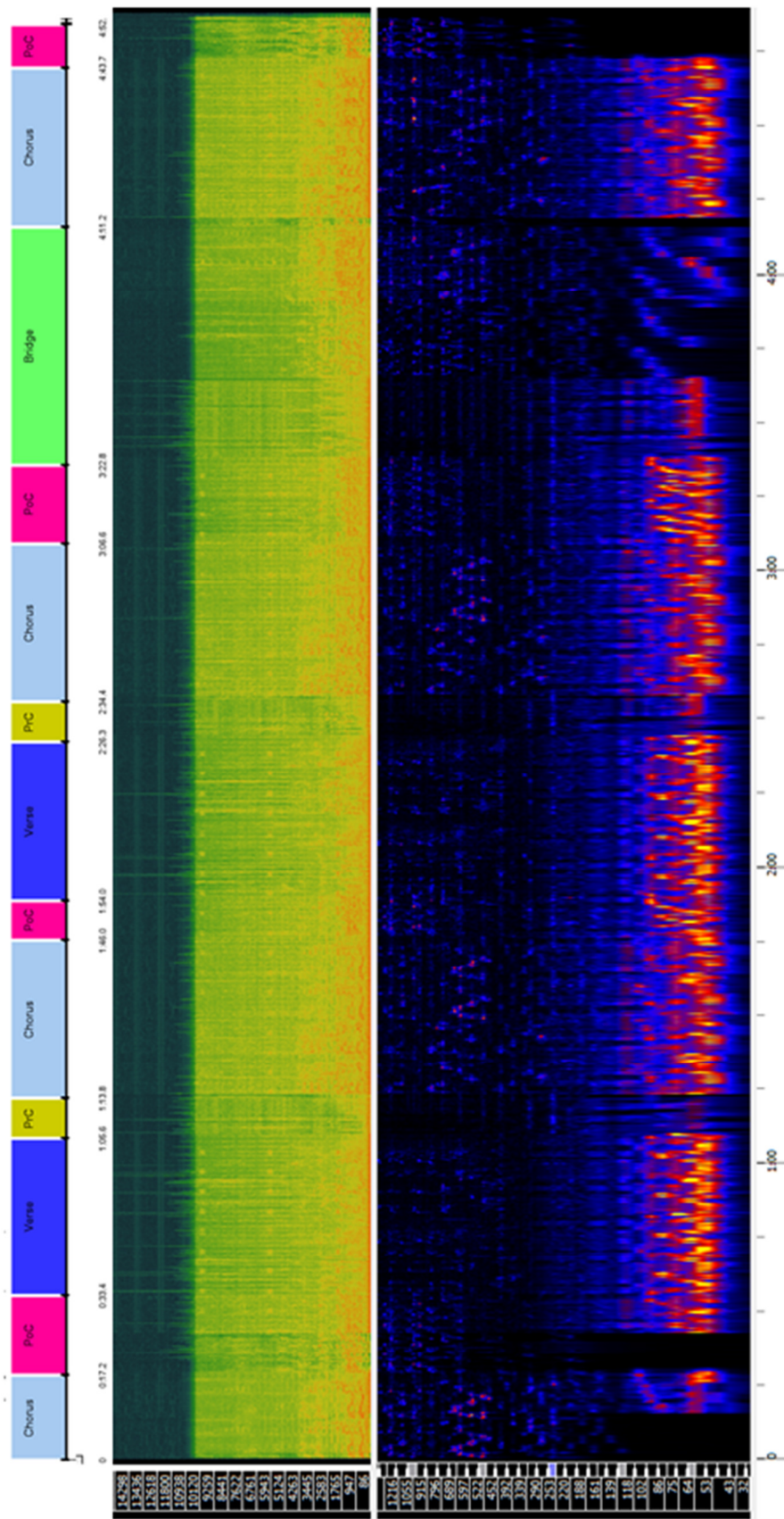
Bad Romance

To the best of my knowledge, Mark Spicer (2011, par. 8-10) is the first to address the post-chorus in a scholarly music theory article.²² As he points out, there are a number of examples of post-choruses in earlier popular music. However, it has become common only recently. In his analysis of Lady Gaga's "Bad Romance" (2009), presented as a video consisting of a timeline with a formal breakdown of the song's sections and accompanying commentary, Spicer points to the post-chorus as a section that consistently follows the different iterations of the chorus (Example 5.15 shows Spicer's analysis of the formal sections positioned above a spectrogram). According to the traditional aspects of the song, this is an uncontroversial

²² A Google search led me to references to the post-chorus in online forums dating back as far as 2006.

interpretation. However, Spicer himself reveals an internal debate, writing: “It establishes the song’s groove and paves the way for the entrance of the VERSE, but I’m calling it a POSTCHORUS.”

By using the word “but,” Spicer acknowledges a possible dilemma – while the term “post-chorus” implies a relationship with the chorus, this section in fact behaves more like a prologue to the first verse rather than an epilogue to the preceding chorus. He nevertheless decides to label this section the post-chorus, and proceeds to do the same with the next three returns of the melodic and lyrical materials of this section (1:46-1:54, 3:06-3:23, and 4:43-4:52). Spicer does not comment on the second and third iterations, but does on the final one: “This final postchorus serves as the song’s CODA.” This statement recognizes a difference in musical behavior between sections that are identical in harmony, melody, and lyrics.



Example 5.15. Lady Gaga, "Bad Romance" (2009). Formal analysis by Mark Spicer (2011).

As evident from the spectrograms, each of the four sections identified by Spicer as the post-chorus features a different sonic behavior. Therefore, from a sonic standpoint, they do not all act as post-choruses. In order to clarify the sonic distinction between the different iterations of what Spicer terms the post-chorus and its formal implications, I will use the term “rah-section” to describe them going forward, instead of “post-chorus.” The first rah-section begins with rather low energy in its first half, but ramps it up in its second half, anticipating the energy of the first verse. As Spicer implies in his comment, this rah-section is sonically an intro. In its next iteration (1:46-1:54), the rah-section indeed behaves like a post-chorus, completely maintaining the level of sonic energy of the chorus. The third rah-section is twice as long as the second, and its second half (3:14-3:22) is identical to it. However, in its first half (3:06-3:14), the high-frequency noise track, which in earlier iterations persists from the chorus, drops out. Moreover, after the first chorus, the high-pitched synthesizer chord accompaniment is replaced in the rah-section by descending glissandi in a similar register. These sounds return only in the second half of the third rah-section. The absence of the noise and high-pitched glissandi substantially reduces the high-frequency content in the first half of the third rah-section, and sonically detaches it from the preceding chorus, as evident in the plain spectrogram. Therefore, I interpret this section as a transition, rather than a post chorus. The fourth and final iteration of the rah-section, as Spicer suggests, serves as an outro.²³ Example 5.16 shows my interpretation of the song sections, in light of the above, taking the sonic dimension into consideration.

²³ Spicer uses the term “coda” to describe the behavior of this section. Although I do not dispute that it behaves like a coda, I prefer to use the term “outro” for two reasons. First, it is more consistent with current pop jargon. Second, “outro” is a term that encompasses several types of ending sections (including coda), and unless there is a reason for pointing out the specific type of ending, I prefer to use the more general term.

The above examples provide a more in-depth glimpse into the decision making process in which I engaged when analyzing these songs. The roles of the traditional and sonic dimensions are sometimes clear, but at other times there is a need to decide which one is dominant in determining the character of a section or to consider a mixed character. For example, although the second half of the verse in Katy Perry's "Unconditionally" begins a stepwise buildup, I could not label it a pre-chorus because it repeated the melody from the first half and was clearly written as the consequent of the verse. Conversely, although the melodies of the pre-chorus and chorus in "Till the World Ends" are identical, the great sonic contrast and the change from lyrics to "ohs" distinguish them as separate sections. In the case of "Bad Romance," there are four sections identical in lyrics, melody, and harmony, but sonically contrasting. The most difficult to categorize is the third rah-section, which I call a transition. Some might interpret the temporary high-frequency loss as negligible, and therefore label this section an additional post-chorus. While this is plausible, I hear the dropping out of the noise and synthesizer accompaniment as a deliberate dip in sonic density, and thus decided to acknowledge this sonic detachment and label it a transition. Analysts might choose to emphasize one over the other, but negotiating the relative role of each of the traditional and sonic dimensions is nevertheless an essential part of analyzing this music.

EXPLORING MORE COMPLEX FORMS

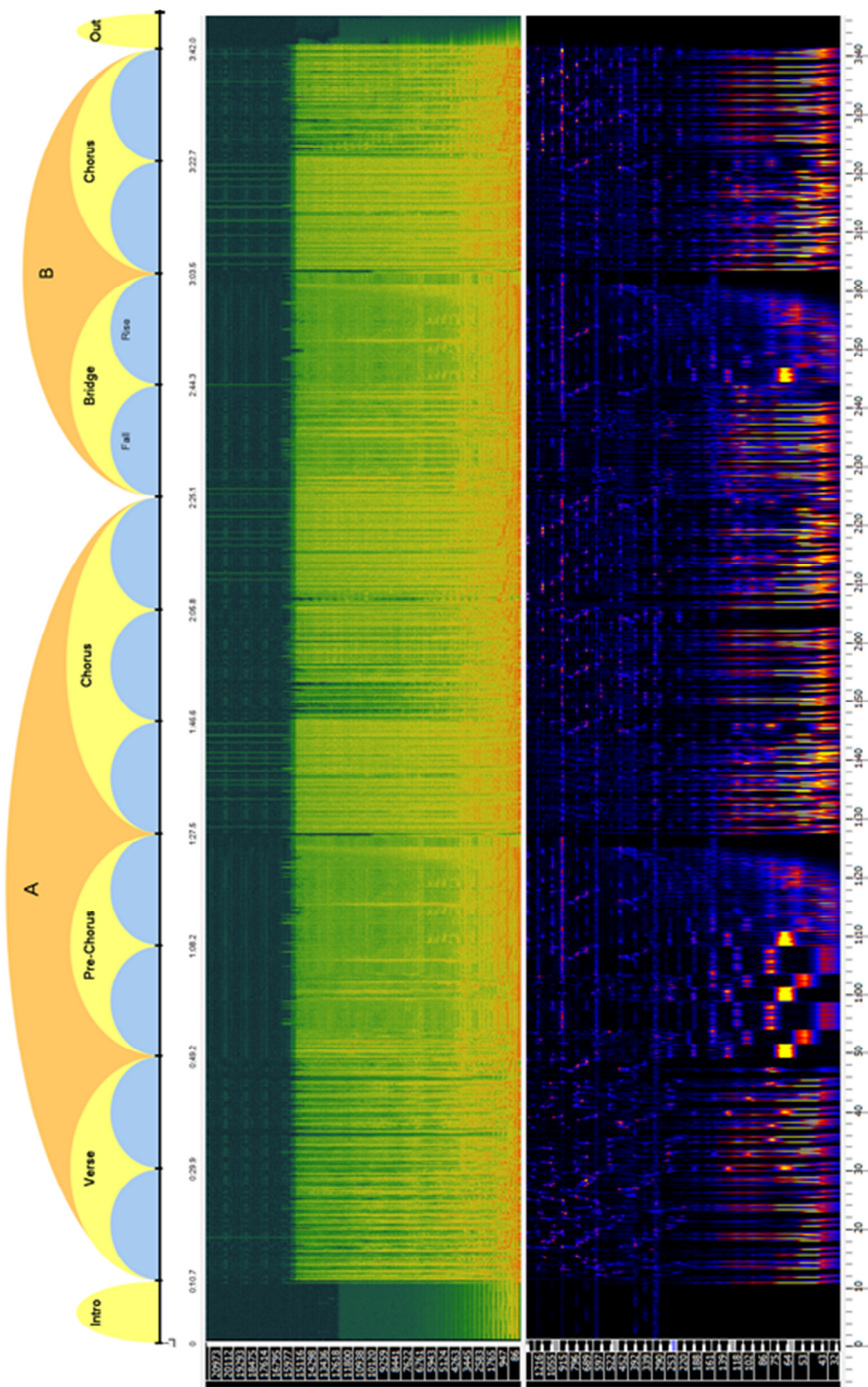
Some pop songs adhere to the simplest structure of verse-chorus form, while others have more complex structures that include optional sections. Most, whether simple or complex, will include the three cycles of setup–buildup–peak that correspond to the traditional verse–pre-

chorus–chorus sequence. In this section, I will explore two of the more analytically challenging songs in my sample. The first – “Roses” (2015) by The Chainsmokers, featuring Rozes – does not follow the typical verse-chorus form, and instead features a more drawn-out structure, in which sections are prolonged and include more internal variation than usual. The second – Rihanna’s “We Found Love” (2011) – is a hybrid pop/EDM song that includes verse-chorus form, but elaborates it in interesting ways. I will discuss the internal debates I experienced while analyzing these songs, in order to illuminate the reasons for my interpretations.

Roses

There are a few things about “Roses” that make a formal analysis of it within the paradigm of verse-chorus form challenging. On the one hand, it has distinct sections that sound like a verse and pre-chorus. On the other hand, while it features a hook (“Say you’ll never let me go”) that is repeated 11 times, it has no section that sounds like a conventional chorus. Moreover, the song never repeats the melody of the opening verse, and the internal variation within sections complicates making decisions about the hierarchy between the traditional and sonic dimensions.

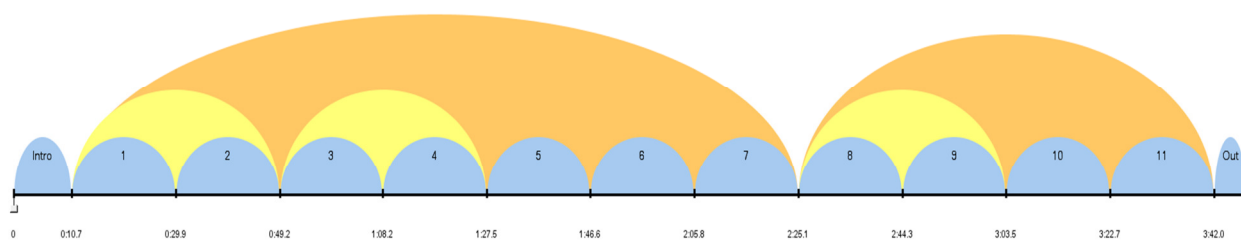
Example 5.17 shows my formal interpretation of “Roses.” On its face, it is a very simple **AB** formal structure. However, though **A** is not repeated, it is drawn out over two minutes and fifteen seconds. Classification of each section, particularly the chorus and the bridge, was not straightforward, due to their melodic and lyrical content. In the following paragraphs, I will discuss the formal possibilities and the reasons for my choices.



Example 5.17. *The Chainsmokers, "Roses" featuring Rozes (2015).*

As I said, “Roses” does not repeat the verse. In fact, the first time any material returns is at 2:25 (the pre-chorus [0:49-1:27] returns as the bridge [2:25-3:03], as shown in Example 5.17), which is very unusual for a pop song. This return of material is the principal structural dividing point. Internally, I heard a notable sonic change every 19 seconds, or every eight measures (assuming $\text{♩} = 100$), but the clearly functional changes were at 0:49, 1:27, 2:25, and 3:03.

My initial mapping is shown in Example 5.18. I have numbered each of the eleven 8-measure sections between 0:10 and 3:42. The intro and outro are quite conventional and easily identified. I could also confidently group sections 1 & 2 (0:10-0:49), 3 & 4 (0:49-1:27), and 8&9 (2:25-3:03), due to the stark sonic contrast between each of them and their surroundings. However, at this point, I had not yet categorized them, because there were unanswered questions about each grouping. I was also not quite sure how to interpret sections 5-7 (1:27-2:05) and 10-11 (3:03-3:42). There is sufficient sonic contrast between these sections to define each as independent. However, when compared with the rest of the song, they could be grouped together, as they share several traits: the harmonic pace, which is different from the rest of the song; identical rhythmic patterns; the absence of lyrics, other than repetitions of the main hook; and similar melodic riffs. Still, I wasn’t quite ready to group them as unified choruses, since it is highly unusual for a pop chorus to exhibit so many internal sonic divisions.



Example 5.18. *The Chainsmokers, “Roses.” Initial formal mapping.*

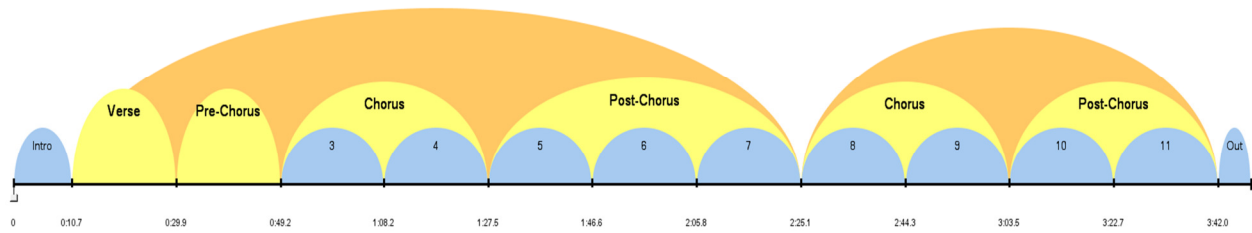
I also had questions regarding the sections that I had already grouped together. For instance, do sections 1 & 2 make up a verse, or is section 2 a pre-chorus? There are two factors supporting the latter interpretation. First, the melody is distinctly different from that of the previous section and contains fragments that sound more like a hook (“We could be beautiful” and “Say you’ll never let me go”). Second, the vocal line is doubled and a bass is added to the mix, revealing the vi-IV-I chord loop that was previously withheld, when only the synth ostinato provided harmonic support to the vocals. Both of these developments are typical, albeit not exclusive, to the pre-chorus.

On the other hand, the more significant sonic buildup takes place in sections 3-4. The overall sonic density increases, as the plain spectrogram shows, but the rhythmic energy drops. The kick drum is eliminated and the bass switches to playing long notes instead of a rhythmic pattern. While there is a new sound (possibly a combined guitar and vocal sample) playing a rhythmic pattern (0:53-0:58 and 1:03-1:08), it does not increase the rhythmic energy, since it comes and goes in this section. This buildup is intensified starting at 1:08, as the bass elevates the rhythmic intensity by playing quarter notes, doubled by the returning kick drum. Gradually, this section turns into a full-fledged riser, with the drums accelerating starting at 1:18, a noise

sweep beginning to ascend at 1:21 as the lower frequency sounds are obscured, and a break occurring at 1:25.

Sonically, this riser in sections 3-4 defines the sections that precede and succeed it. Structurally, sections 1-2 function as the setup, sections 3-4 as the buildup, and the peak arrives at section 5. Moreover, while there are momentary internal dips in sonic density in sections 5-7, the persistence of the kick drum and bass, along with the other shared traits listed above, help to sustain the feeling of peak sonic energy throughout these sections. Therefore, I decided to group sections 5-7 as the sonic peak. The same applies to sections 10-11, which repeat sections 5-7, but do so in an interesting way. While this repetition is only two sections long, no section is in fact omitted. Instead, section 5 is repeated as is in section 10, and sections 6 and 7 are juxtaposed with minor revisions to form section 11.

While at this point I have mapped out the sonic dimension of the song, this does not automatically mean that sections 1-2 make up the verse, sections 3-4 the pre-chorus, and sections 5-7 the chorus. Earlier in this chapter, I showed examples of EDM-oriented songs in which the buildup occurs in the chorus, and is followed by a dance breakdown with an aggressive synthesizer riff in the foreground. The sonic sequence in “Roses” certainly resembles this kind of structure. Following this model, section 1 could be the verse, section 2 the pre-chorus, sections 3-4 the chorus, and sections 5-7 the post-chorus, as shown in Example 5.19.



Example 5.19. *The Chainsmokers, “Roses.” Rejected alternative formal mapping.*

Ultimately, I rejected this interpretation, for two reasons. First, sections 5-7 in “Roses” contain the vocal hook, which is not typical of EDM-oriented post-choruses such as those in “Heroes” and “Outside.” Sections that heavily feature the vocal hook usually constitute the chorus, as evident by songs such as Jason Derulo’s “Wiggle” (2014) and Jack Ü’s “Where are Ü Now” (2015). Second, in “Heroes” and “Outside” the vocal hooks were present throughout the buildups, and from the perspective of the traditional dimension these sections were unquestionably the chorus. This is not the case in sections 3-4 of *Roses*, in which neither the sonic dimension nor the traditional dimension supports categorizing it a chorus, particularly section 3. By any interpretation, sections 3-4 comprise a pre-chorus and not a chorus. Therefore, sections 1-2 function as the verse, and sections 5-7 make up a 3-part breakdown which functions as the chorus.

Sections 10-11, then, make up the second (and final) chorus, since they repeat sections 5-7 as described above. But what about sections 8-9? From the traditional perspective, they are surely a pre-chorus, as they constitute a return of the melody and lyrics of sections 3-4. However, here the sonic perspective takes precedence, due to the significance of the sonic event at 2:25. The removal of all sounds except for the vocals and kick drum marks the sharpest fall in sonic density in the entire song. As previously mentioned, this is the most idiomatic signal that the

bridge has arrived, and indeed sections 8-9 behave like a bridge. The fall occurs in section 8, while section 9 acts as the rise. Since the bridge marks the arrival of the **B** section, subsections 1-7 therefore make up the **A** section, encompassing a verse, pre-chorus, and chorus, while sections 8-11, containing the bridge and final chorus, comprise the **B** section.

“Roses” exemplifies a more complex relationship between the traditional and sonic dimension than in earlier examples. Although a purely sonic analysis could have led me to the interpretation in Example 5.19, taking the harmony, lyrics, and melody into account ultimately led me to a different conclusion. Conversely, had I relied exclusively on a traditional approach, I would have defined sections 8-9 as a pre-chorus rather than a bridge. Doing so would have changed my understanding of the large-scale structure of this song. Since my point of reference was the **AA'B** verse-chorus form, categorizing sections 8-9 as a pre-chorus would have led me to view the entire song as a drawn out **AA'** form, with **B** omitted, rather than **AB**, with **A'** omitted. Taking both dimensions into account, the latter is, in my opinion, a more accurate representation of the song.

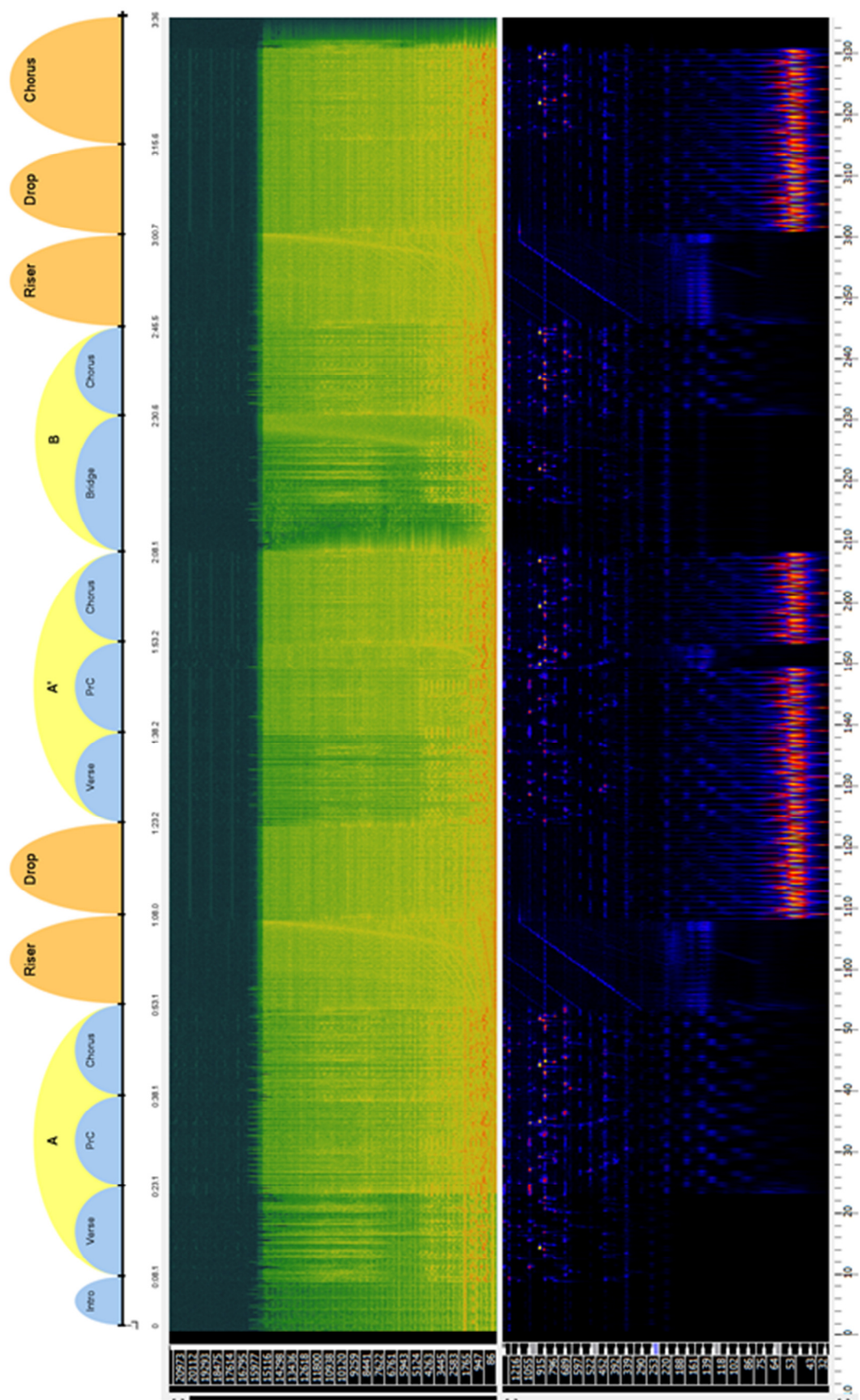
We Found Love

I was presented with a different challenge when I first approached Rihanna’s “We Found Love.” While it clearly contains all the components of a song in **AA'B** verse-chorus form, it features additional sections that not only have significant sonic presence, but also seem to be central to the song’s large-scale structure, rather than external transitions. Representing these sections as external to the song’s core structure seemed to miss its point. For instance, between the start of the first verse and the start of the second verse there are five clearly distinct sections. The first three are easy to categorize as the verse, pre-chorus, and chorus (0:08-0:53). However,

these are followed by a huge riser and a highly-climactic dance section (0:53-1:23), in the mold of the dance post-chorus discussed earlier in this chapter. The song then continues in a conventional manner to an additional verse – pre-chorus – chorus cycle (1:23-2:08), followed by a bridge and a chorus (2:08-2:45), completing all of the obligatory portions of the **AA'B** verse-chorus form. This chorus is succeeded by another iteration of the riser, followed by another climactic dance section, which is extended by another return of the chorus.

Example 5.20 shows this initial mapping. The blue bubbles represent the components of the verse-chorus form, grouped by the larger yellow bubbles, indicating the internal **AA'B** structure. The brown bubbles represent the additional sections that are seemingly outside of the verse-chorus form but make up the most sonically dramatic portions of the song. The drops (1:08-1:23 and 3:00-3:15) are the dramatic climaxes toward which the risers (0:53-1:08 and 2:45-3:00) significantly elevate the feeling of anticipation. How can the dramatic trajectory of this song be traced in a continuous manner that incorporates all of these sections?

When self-contained, each of the **A**, **A'**, and **B** large sections in Example 5.20 have a sonic setup, buildup, and peak. In **A** and **A'**, these sonic functions correspond to the verse, pre-chorus and chorus. In **B**, the setup and buildup occur in the bridge, and the peak in the chorus. However, similarly to section 2 in “Roses” – which locally acted as a stepwise buildup, but was ultimately grouped as the second half of a deeper-level structural setup – the risers and climactic dance sections in “We Found Love” sonically overshadow the sonic developments within the **A**, **A'**, and **B** sections, and define each of these sections as a unified, deeper-level sonic function.



Example 5.20. Rihanna, "We Found Love" (2011). Initial formal mapping.

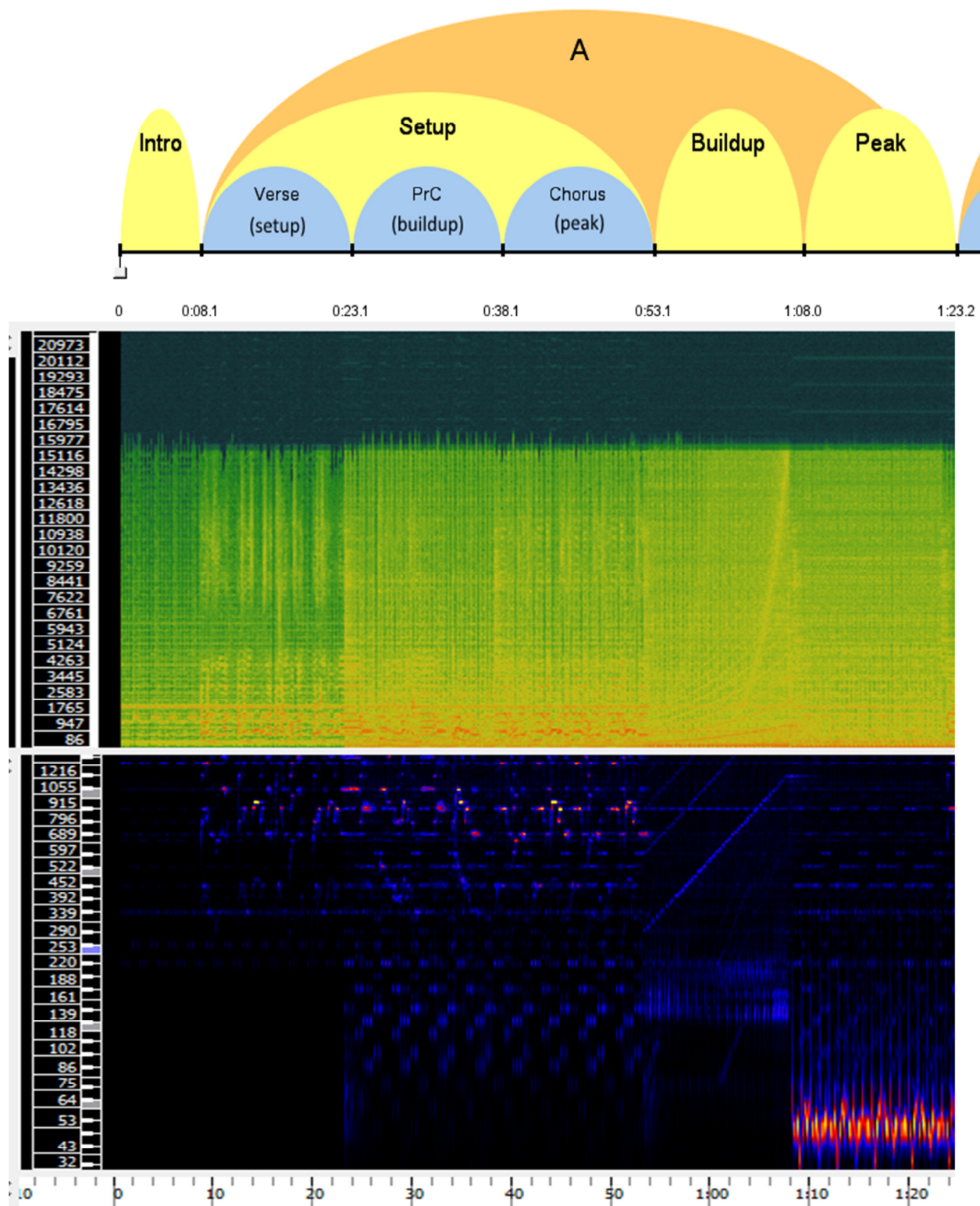
In order to understand this, let us look at the first five sections that I mentioned earlier. The verse (0:08-0:23) features only a keyboard riff and an almost-inaudible percussive sound playing perpetual sixteenth notes as the accompaniment to Rihanna's voice, setting the baseline sonic energy for the song. The pre-chorus (0:23-0:38), fulfilling the buildup function, starts with a cymbal "splash" and adds significant sonic layers, including a bass and rhythmic claps that increase the sonic energy. The chorus (0:38-0:53), while also beginning with a cymbal splash, adds only an extremely quiet noise track in the background. Otherwise, it is mostly a "peak by association" based on its melodic and lyrical content - it contains the hook ("We found love in a hopeless place"), which combines with a returning melody from the verse to create an anthem-like chorus.

This setup-buildup-peak sequence is extremely subdued compared to other songs in this genre. Aside from the addition of the vocals, there is no sonic change between the intro and the verse, and as mentioned, this is the case between the pre-chorus and the chorus as well. The only significant increase in sonic energy takes place at the start of the pre-chorus, but even this intensification is rather mild. The kick drum has yet to be introduced and the bass is playing in a relatively high, borderline-mid-frequency register. There is also little sonic presence in the high-frequency range. The texture is thus relatively focused in the mid-frequency region. It is unusual in a dance-oriented song to have so little sonic intensification between the intro and the chorus, but this hints at how the riser-drop sequences relate to the sections that comprise the **AA'B** verse-chorus form.

Although such restrained sonic development may at first sound underwhelming to the listener (especially in the dance club), it has a desirable effect – the moderate sonic energy of the chorus leaves the song room to intensify. A verse–pre-chorus–chorus sequence such as those in

“Till the World Ends” or “I Knew You Were Trouble” would substantially reduce the effectiveness of the additional buildup-peak progression that follows. The chorus in these songs already represents such a high level of sonic energy, that even if the next peak was more climactic, the range of sonic intensification would be quite small. The relatively tamed sonic development in “We Found Love” leading up to 0:53 allows the following riser and drop to make a significant sonic impact. In effect, the riser serves as an instrumental pre-chorus, fulfilling the buildup function, and transforms the entire preceding sequence into a super-verse, which functions as a prolonged setup. The section following the drop, then, acts as an instrumental chorus, and is the sonic peak. This forms a complex structure, in which there are two structural levels, as shown in Example 5.21. On one level, the verse, pre-chorus, and chorus form their own setup-buildup-peak sequence. On a deeper structural level, they are contained within a setup, which proceeds to a more substantial buildup and peak.

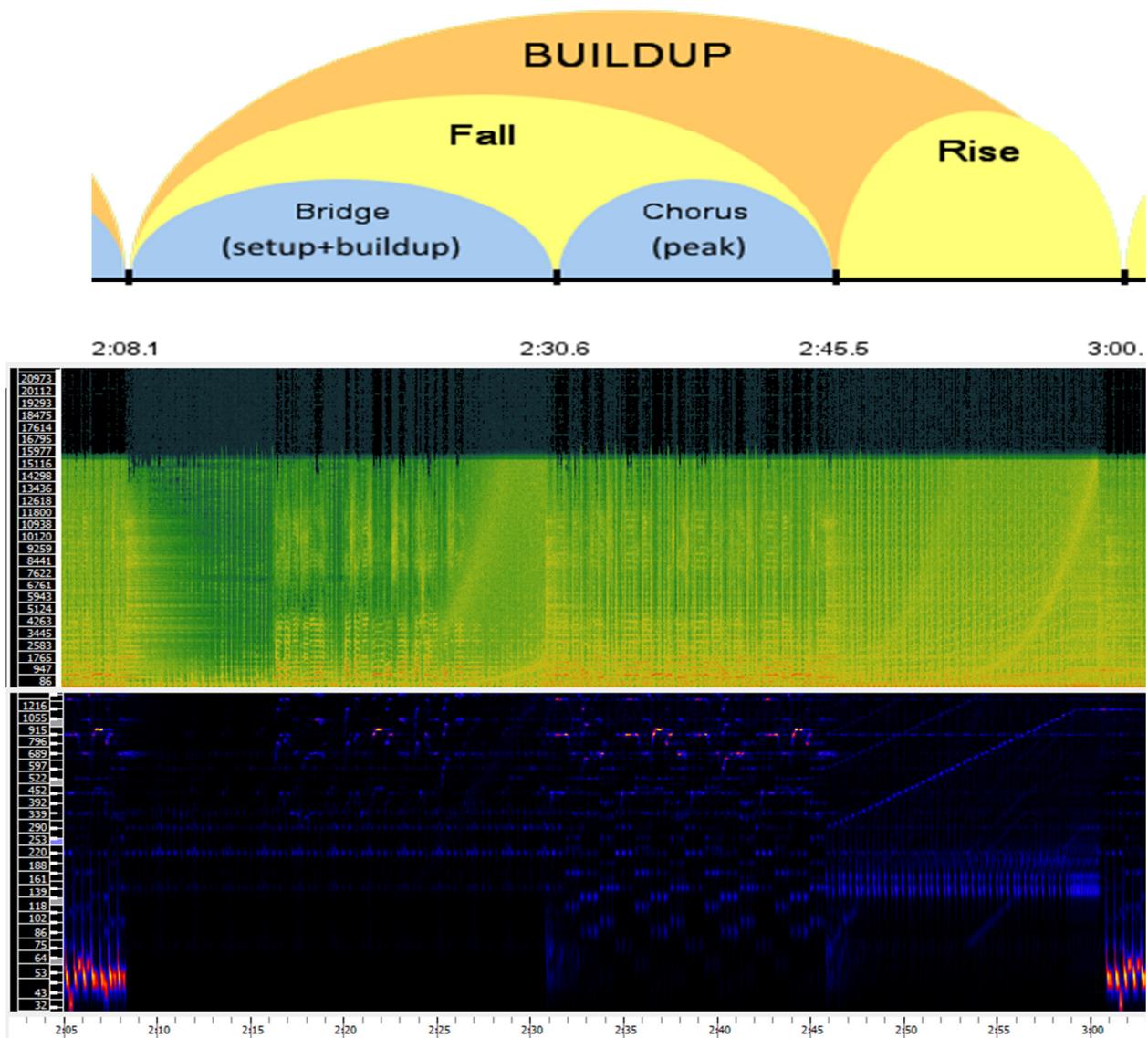
This sonic connection between the sections that comprise the large A-section hints at what is to come later in the song. All of the sections discussed above are repeated, but in different contexts. First, the verse, pre-chorus, and chorus return immediately following the end of the peak section (1:23-2:08). This sequence differs from its previous iteration in two ways. First, it starts at a much higher level of sonic energy. Second, the energy is further elevated at the start of each new section, causing the verse, pre-chorus, and chorus to form a more substantial setup-buildup-peak sequence. Yet, because the second verse started at a relatively high energy level, the range of intensification is overshadowed by the wide sonic range of the previously heard riser, a fact that affects the overall role of this sequence in the song’s sonic trajectory.



Example 5.21. *Rihanna, "We Found Love," 0:00-1:23.*

The bridge (2:08-2:30) is not divided into a fall and a rise, fulfilling both functions by initially dropping to the lowest level of sonic activity and ending with a pronounced noise sweep. The chorus that follows (2:30-2:45) is the final section of the internal verse-chorus structure. If these sections constituted a standalone song, this chorus would be expected to be the most sonically energetic. Yet, instead of being at least as climactic as the second chorus, it is as subdued as the first. The reason for this becomes apparent once we hear the big riser return (2:45-3:00). As we have seen in this song and in previous examples, the most pronounced buildups define their surroundings. In this case, the preceding bridge and chorus had a setup, buildup (both in the bridge), and peak (chorus), but on a deeper structural level, they act as a setup to the big riser.

When zooming out to an even-deeper structural level, we hear that the entire sequence between 2:08 and 3:00 is a large-scale buildup with an internal fall-rise, or in other words, a “super-bridge,” as shown in Example 5.22. Since, as mentioned, the beginning of the bridge represents the lowest point of sonic energy in the song, the range of intensification within this time frame is the largest in the song, and therefore these three sections (the bridge, third chorus, and riser) comprise the most substantial structural buildup.

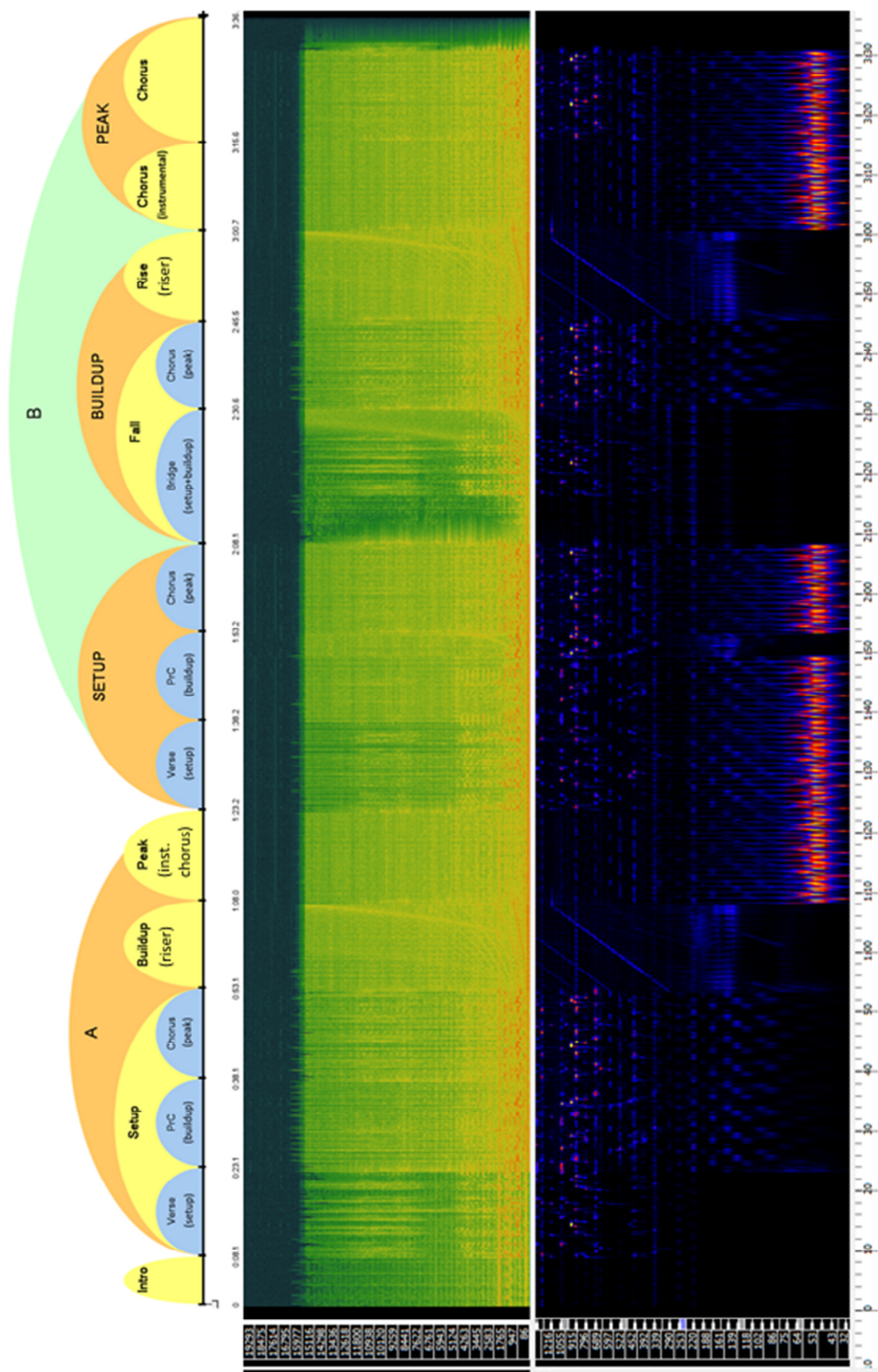


Example 5.22. *Rihanna, “We Found Love,” 2:08-3:00.*

While it is clear that this large-scale buildup results in a climactic peak, it also defines the verse, pre-chorus, and chorus that preceded the bridge as an additional super-verse, or a large-scale setup. This completes a mega-structure that includes a setup (1:23-2:08), buildup (2:08-3:00), and peak (3:00-3:36). Within this structure, the returning rise(r) and peak sections (both

identical to their previous iteration) each becomes a component in an expanded buildup and peak, respectively. The riser combines with the bridge and third chorus to form a prolonged buildup, and the instrumental chorus is followed by a vocal chorus, extending and intensifying the final peak section.

In essence, there is a “song within a song” in “We Found Love.” The relationships between the components of the **AA’B** verse-chorus structure and the EDM-oriented risers and climactic sections create an interesting new structure, as mapped in Example 5.23. The first five sections of the song (excluding the intro) comprise a single large section in which the setup-buildup-peak sequence operates on two different structural levels (represented by the blue and yellow bubbles). However, this large section turns out to be a smaller version of the rest of the song, in which the sonic sequence operates on three structural levels (represented by the blue, yellow, and brown bubbles). These relationships are made possible by limiting the sonic range in the “internal song.” The first chorus and the sections leading up to it are subdued in order to allow the riser and instrumental chorus to function as a more substantial buildup-peak progression; the range of intensification between the second verse and the second chorus is limited, allowing it to act as a large-scale setup; and the bridge and third chorus are also limited in their range of intensification, which enables them to be a setup for the riser while combining with it to form a deeper-level buildup.



Example 5.23. Rihanna, "We Found Love." Final formal analysis.

“We Found Love” stretches the traditional/sonic relationship beyond the normal divergence of dimensions heard in songs like “I Knew You Were Trouble” or “Unconditionally.” It creates a much more complex interaction between the two dimensions of the form. Instead of progressing on parallel routes, one dimension is completely embedded within the other. While there are other songs that follow unusual and complex forms, this “song within a song” structure is unique to this song, to the best of my knowledge.

CONCLUSION

The phenomena described in this chapter are what make contemporary pop music unique. Sound has been a primary concern for pop and rock record producers since the mid-20th century, but that it is a consistent leading factor in shaping the syntax and form of popular songs is a new development. In electronic dance music, sound is indeed in the foreground of dramatic development, but the degree to which the traditional and sonic dimensions in today’s pop are intertwined in determining a song’s form makes it a completely different animal from other popular genres, and this relationship must be taken into consideration when analyzing this music.

CHAPTER 6

CONCLUSION

The primary aim of my research is to illuminate a sonic syntax that is central to contemporary pop music, and which relies on timbre, gesture, and spatialization. To review, Chapters 3, 4, and 5 propose a framework by which the components of this syntax can be categorized and analyzed. The sonic syntax largely consists of the three major functions that I propose – *setup*, *buildup*, and *peak*. These functions progress in cycles on multiple structural levels. The setup and peak represent relative low and high points of energy in a sonic cycle. Between them, the buildup generates anticipation toward the peak by using a range of techniques, which include adding sonic layers at the beginnings of subsections, abruptly removing sonic layers to drop the sonic energy below the song's baseline level, and employing ascending linear gestures such as filter sweeps and pitch bends (*glissandi*).

The emergence of this syntax in contemporary pop music is a result of the increasing influence of electronic dance music (EDM). As sonic manipulations became increasingly dominant in generating syntax in pop songs in the early 21st century, tonal behavior became less syntactic. The recurring chord loops in this genre often feature chord combinations that do not

adhere to the traditional cycle of harmonic functions. Anticipatory mechanisms such as dissonant sonorities and chromaticism have nearly disappeared from this music. In today's pop, anticipation and form are defined in large part by their sonic treatment.

Importance of this Research

In Chapter 1, I quoted David Blake: "...timbre is especially frustrating for analytic description, at once the most apparent and least systematizable musical parameter." Although nearly every genre is distinguished from others by its timbre, it is difficult to categorize and analyze its syntactical properties. Twenty first century pop music, however, uses timbre and other sonic features in a highly systematized and categorizable way, providing an opportunity to develop an analytical framework that incorporates these musical aspects. This framework can potentially serve as a point of departure for developing new models for examining timbre in other genres, even in those where it operates as a subordinate musical feature.

My initial motivation for pursuing this music as a research topic, aside from my personal fondness for it, was my view that music that is so culturally impactful should be addressed from a music theoretical perspective. Moreover, although there are clear advantages to investigating musical genres for which a historical perspective exists, I found the prospect of exploring a genre as it evolves quite intriguing.

This research also stresses the importance of technology as a means of creating music, as well as for analyzing it, in the 21st century. As technological advancements are made at an increasingly rapid pace, more and more music is created by using advanced audio and synthesis software. Music theorists who wish to explore this music cannot rely solely on their musical

training. They must gain technological proficiency, which includes the ability to use digital audio workstations, read spectrograms, and develop aural skills that address timbre and spatialization.

Applicability for Practitioners

As part of my research, I have engaged in conversations with producers, songwriters, audio engineers, artists, and record label executives. One of the ways by which I tested the applicability of my research to pop practitioners was to analyze works in progress given to me by producers and songwriters, and to offer them feedback based upon the framework that I have developed. My purpose was not to provide them with a prescriptive formula for writing a top-40 hit, but to illuminate the effects of certain artistic decisions in a way that will enhance their ability to critically examine their own work. Topics of conversation included the effects of local decisions on the overall sonic trajectory of a song. For instance, the impact of a drop in sonic energy in the second verse on the effectiveness of a similar drop in the bridge was a focal point of one conversation. In another, a songwriter's use of a closed chord loop (e.g., I-vi-V, looping back to I) seemed to have an inhibiting effect on some sonic processes, whereas experimenting with open chord loops that omit either V or I allowed these processes to become more effective. Studying these works-in-progress and engaging in these conversations has helped to shape some of the directions of inquiry I pursued, with the intention of making my research a useful tool for these practitioners.

Final Thoughts

"It keeps changing all the time. We've just made it out from the marshlands of EDM. Nothing wrong about EDM, great songs came out of it, but there was a period when everything had to have a pace of ha 128 bpm and be DJ-related. These days, there's no dominating trend among the Top 40-songs, and I really enjoy that. A hit can be someone

just singing to piano music, anything. But back to your question: I recently re-watched an old movie that I used to like when it came out. Now that I watched it over, I felt the movie's tempo. It all felt a bit slow. They showed the whole trip to the airport. Today it's more 'Boom!' and you're at the airport. The same thing has happened to pop music. There's less downtime. Pop music follows the evolution of society in general. Everything moves faster. Intros have gotten shorter.”²⁴

Max Martin, 2016

Contemporary pop music is often disparaged in social media and online forums. A perceived lack of musicianship on the part of the artists and creators of this music is often cited as a reason for its inferiority in comparison to other popular music genres, jazz, and concert music. This perception is, in my opinion, a result of this music being judged based on parameters that are far less relevant to it than they are to other genres. Using an unchanging chord loop through an entire song may be perceived as a result of poor musical training exhibited by the song's composer, but in fact it allows the song's sonic syntax to be foregrounded and independent. Vocal processing is often the subject of criticism of an artist's musical talent, but producers often use it as a means of creating new timbres that are impossible to achieve otherwise. The purpose of this research is not to suggest that every recently released pop song is a masterpiece. My hope is that it provides a perspective that allows the listener to judge this music on its own terms, rather than to rely on principles that apply to other genres.

²⁴ Jan Gradvall, “World Exclusive: Max Martin, #1 Hitmaker.” *Dagens industri Långläsning*. 2016.

APPENDIX – ANALYZED SAMPLE OF POP SONGS 2008-2016

| ARTIST | SONG TITLE | YEAR RELEASED |
|--|-------------------------------|--------------------------|
| Alessia Cara | Here | 2015 |
| Alesso featuring Tove Lo | Heroes | 2014 |
| Ariana Grande | Break Free | 2014 |
| | Dangerous Woman | 2016 |
| | Focus | 2016 |
| | One Last Time | 2014 |
| | Right There | 2013 |
| Ariana Grande featuring Iggy Azalea | Problem | 2014 |
| Ariana Grande featuring Mac Miller | The Way | 2013 |
| Ariana Grande featuring Nathan Sykes | Almost is Never Enough | 2013 |
| Ariana Grande featuring The Weeknd | Love Me Harder | 2014 |
| Becky G. | Break a Sweat | 2015 |
| | Can't Stop Dancing | 2014 |
| | Lovin' So Hard | 2015 |
| | Shower | 2014 |
| | 3 | 2009 |
| Britney Spears | Hold it Against Me | 2011 |
| | I Wanna Go | 2011 |
| | Till the World Ends | 2011 |
| | Outside | 2014 |
| Calvin Harris featuring Ellie Goulding | Call Me Maybe | 2012 |
| Carly Rae Jepsen | Roses | 2016 |
| The Chainsmokers featuring Rosez | Your Body | 2012 |
| Christina Aguilera | Body Party | 2013 |
| Ciara | Get Lucky | 2013 |
| Daft Punk | Lose Yourself to Dance | 2013 |
| | Hey Mama | 2014 |
| David Guetta featuring Nicki Minaj | | |

| | | |
|-------------------------------------|-------------------------------------|-------------|
| Daya | Hide Away | 2015 |
| Demi Lovato | Cool for the Summer | 2015 |
| DJ Snake featuring AlunaGeorge | You Know You Like It | 2013 |
| Drake featuring Rihanna | Take Care | 2011 |
| Ellie Goulding | Burn | 2013 |
| | Lights | 2011 |
| Fifth Harmony featuring Kid Ink | Worth it | 2015 |
| Gotye featuring Kimbra | Somebody that I Used to Know | 2011 |
| Hozier | Take Me To Church | 2013 |
| Icona Pop | All Night | 2013 |
| Icona Pop featuring Charlie XCX | I Love It | 2013 |
| Jason Derulo | Want to Want Me | 2015 |
| | Whatcha Say | 2010 |
| Justin Bieber | Boyfriend | 2012 |
| | Sorry | 2015 |
| | What do You Mean | 2015 |
| Justin Bieber featuring Nicki Minaj | Beauty and a Beat | 2012 |
| Katy Perry | E.T. | 2010 |
| | I Kissed a Girl | 2008 |
| | Roar | 2013 |
| | Teenage Dream | 2010 |
| | The One that Got Away | 2010 |
| | Unconditionally | 2013 |
| Katy Perry featuring Snoop Dog | California Gurls | 2010 |
| Kelly Roland featuring Lil Wayne | Motivation | 2010 |
| Kendrick Lamar | Bitch, Don't Kill My Vibe | 2012 |
| | Swimming Pools | 2012 |
| Lady Gaga | Bad Romance | 2009 |
| Lana Del Rey | Summertime Sadness | 2012 |
| Lil Wayne featuring Drake | She Will | 2011 |
| Lorde | Royals | 2013 |

| | | |
|---|---------------------------------|-------------|
| Ludacris featuring Kelly Rowland | Representin' | 2015 |
| Major Lazer and DJ Snake | Lean On | 2015 |
| Miguel | Adorn | 2012 |
| | How Many Drinks | 2012 |
| Miley Cyrus | Adore You | 2013 |
| | See You Again | 2008 |
| | We Can't Stop | 2013 |
| | Wrecking Ball | 2013 |
| Nicky Minaj | The Night is Still Young | 2015 |
| | Starships | 2010 |
| Pitbull, Ne-Yo | Time of Our Lives | 2014 |
| Rihanna | American Oxygen | 2015 |
| | Pour it Up | 2012 |
| Rihanna featuring Calvin Harris | We Found Love | 2011 |
| | Work | 2016 |
| Rihanna featuring Mikky Ekko | Stay | 2012 |
| Robin Thicke featuring Nicky Minaj | Back Together | 2015 |
| Sebastian Mikael featuring Wale | Last Night | 2014 |
| Selena Gomez | Good for You | 2015 |
| | Same Old Love | 2015 |
| Sia | Chandelier | 2014 |
| | Elastic Heart | 2014 |
| Skrillex and Diplo (Jack Ü) featuring Justin Bieber | Where are Ü Now | 2015 |
| Taio Cruz | Dynamite | 2010 |
| Taylor Swift | Blank Space | 2014 |
| | I Knew You Were Trouble | 2012 |
| | Out of the Woods | 2015 |
| | Shake it Off | 2015 |
| | Style | 2015 |

| | | |
|------------------------------------|--|-------------|
| Taylor Swift (<i>cont.</i>) | We Are Never Ever Getting Back Together | 2012 |
| | Wildest Dreams | 2015 |
| | You Belong With Me | 2008 |
| Tove Lo | Habits | 2014 |
| | Not on Drugs | 2014 |
| | Talking Body | 2014 |
| | Timebomb | 2014 |
| Usher | DJ Got Us Falling In Love Again | 2010 |
| | More | 2010 |
| The Weeknd | Earned It | 2015 |
| | Wicked Games | 2011 |
| Will.I.Am featuring Britney Spears | Scream and Shout | 2013 |
| Zayn | Pillow Talk | 2016 |

BIBLIOGRAPHY

- Abbate, Carolyn. "What the Sorcerer Said." *19th-Century Music* 2:3 (Spring 1989), pp. 221-230.
- Adams, Kyle. "What Did Danger Mouse Do? *The Grey Album* and Musical Composition in Configurable Culture." *Music Theory Spectrum* 37:2 (Fall 2015), pp. 7-24.
- Almén, Byron. "Narrative Archetypes: A Critique, Theory, and Method of Narrative Analysis." *Journal of Music Theory* 47:1 (Spring 2003), pp. 1-39.
- Attas, Robin. "Form as Process: The Buildup Introduction in Popular Music." *Music Theory Spectrum* 37:2 (Fall 2015), pp. 275-296.
- Barber, Simon. "Soundstream: The Introduction of Commercial Digital Recording in the United States." *Journal on the Art of Record Production*. November 2012.
- Bates, Eliot. "What Studios do." *Journal on the Art of Record Production*. November 2012.
- Benjamin, Walter. "The Work of Art in the Age of Mechanical Production." *Illuminations*. London, Pimlico 1999.
- Berger, Karol. "Diegesis and Mimesis: The Poetic Modes and the Matter of the Artistic Presentation." *The Journal of Musicology*. 12:4 (Autumn 1994), pp. 407-433.
- Blake, David K. "Timbre as Differentiation in Indie Music." *Music Theory Online*. 18:2. June 2012.
- Boone, Christine. "Mashing: Toward a Typology of Recycled Music." *Music Theory Online* 19:3. September 2013.
- Burnham, Scott. *Beethoven Hero*. Princeton: Princeton University Press. 2000.
- Butler, Mark J. "Hearing Kaleidoscopes: Embedded Grouping Dissonance in Electronic Dance Music." *Twentieth-Century Music* 2:2. United Kingdom: Cambridge University Press. September 2005, pp. 221-243.
- _____. "Turning the Beat Around: Reinterpretation, Metrical Dissonance, and Asymmetry in Electronic Dance Music." *Music Theory Online* 7:6. December 2001.
- Campelo, Isabel. "From LA to Lisbon: The "LA Sound" as a Referential Production Sound in Rui Veloso's Recording Career." *Journal on the Art of Record Production*. June 2012.

- Cannam, Chris, Christian Landone, and Mark Sandler. "Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files." *Proceedings of the ACM Multimedia 2010 International Conference*. October 2010, pp. 1467-1468.
- Caplin, William E. *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*. New York, Oxford University Press, 1998.
- Covach, John. 2005. "Form in Rock Music: A Primer." In *Engaging Music: Essays in Music Analysis*, ed. Deborah Stein, 65–76. New York: Oxford University Press.
- Crooks, John. "Recreating an Unreal Reality: Performance Practice, Recording, and the Jazz Rhythm Section." *Journal on the Art of Record Production*. June 2012.
- Davis, Robert. "Modes of Production, Modes of Listening: Alternative Realities and the Sonic Divide." *Journal on the Art of Record Production*. July 2011.
- De Bono, Edward, *Lateral Thinking – Creativity Step by Step*. New York: Harper & Row, 1970.
- Draper, Paul and Stephen Emmerson. "Remixing Modernism: Re-Imagining the Music of Berg, Schoenberg and Bartok in our Time." *Journal on the Art of Record Production*. July 2011.
- Endrinal, Christopher. "Burning Bridges: Defining the Interverse in the Music of U2." MTO 17.3: Endrinal, Defining the Interverse. October 1, 2011. Accessed February 8, 2015. <http://www.mtosmt.org/issues/mto.11.17.3/mto.11.17.3.endrinal.html>.
- Everett, Walter. "Confessions from Blueberry Hell, or, Pitch Can Be a Sticky Substance." In *Expression in Pop-Rock Music: A Collection of Critical and Analytical Essays*. New York: Garland, 2000, pp. 269-345.
- _____. "Making Sense of Rock's Tonal Systems." *Music Theory Online* 10:4. December 2004.
- Farnell, Andy. *Designing Sound*, Cambridge, MIT Press, 2010.
- Goodchild, Meghan. "Acoustic and musical features of emotional response to orchestral gestures." *Proceedings*, ICA 2013, Montreal.
- Gradvall, Jan. "World Exclusive: Max Martin, #1 Hitmaker." *Dagens industri Långläsning*. February 2016. Web. 13 February 2016.

- Hajimichael, Mike. "Virtual Oasis - Thoughts and Experiences about Online Based Music Production and Collaborative Writing Techniques." *Journal on the Art of Record Production*. July 2011.
- Heiser, Marshall. "SMiLE: Brian Wilson's Musical Mosaic." *Journal on the Art of Record Production*. November 2012.
- Howlett, Mike. "The Record Producer as Nexus." *Journal on the Art of Record Production*. June 2012.
- Huber, David Miles and Robert E. Runstein. *Modern Recording Techniques*, Oxford, Elsevier Inc., 2005.
- Huron, D. *Sweet Anticipation: Music and the Psychology of Expectation*, Cambridge, Bradford Books, 2006.
- Inglis, Sam. "Creative Mix Automation In Your DAW." *Soundonsound*. August 2011. Web. 04 January 2015.
- Katz, Mark and Samantha Bennett, "Technology, Time and Place." *Journal on the Art of Record Production*. November 2012.
- Knowles, Julian and Donna Hewitt. "Performance Recordivity: Studio Music in a Live Context." *Journal on the Art of Record Production*. June 2012.
- Liu-Rosenbaum, Aaron. "The Meaning in the Mix: Tracing a Sonic Narrative in 'When the Levee Breaks'." *Journal on the Art of Record Production*. November 2012.
- Marrington, Mark. "Experiencing Musical Composition in the DAW: The Software Interface as Mediator of the Musical Idea." *Journal on the Art of Record Production*. July 2011.
- McAdams, Stephen. "Perspectives on the Contribution of Timbre to Musical Structure." *Computer Music Journal*, 23:3(Autumn 1999). pp. 85–102.
- _____. "Timbre as a structuring force." *Proceedings*, ICA 2013, Montreal.
- McClary, Susan, and Robert Walser. "Start Making Sense! Musicology Wrestles with Rock." In *On Record: Rock, Pop, and the Written Word*, ed. Simon Frith and Andrew Goodwin. New York: Pantheon Books, 1990, pp. 277-292.
- Mishra, Jyoti. "Akira Kiteshi: From Dubstep to Disney." *Soundonsound*. April 2012. Web. 13 May 2013.
- Moore, Allan F. and Ruth Dockwray. "Configuring the Sound-Box 1965-1972." *Popular Music*. 29:2 (2010), pp. 181-197.

- Moore, Allan. "The So-Called 'Flattened Seventh' in Rock." *Popular Music*. 14:2 (May 1995), pp. 185-201.
- Moore, Austin. "All Buttons In: An Investigation into the Use of the 1179 FET Compressor in Popular Music Production." *Journal on the Art of Record Production*. June 2012.
- Moorefield, Virgil. *The Producer as Composer, Shaping the Sound of Popular Music*, Cambridge, MIT Press, 2005.
- Morey, Justin and Phillip McIntyre. "'Working Out the Split': Creative Collaboration and Assignment of Copyright Across Differing Musical Worlds." *Journal on the Art of Record Production*. July 2011.
- Mynett, Mark. "Achieving Intelligibility Whilst Maintaining Heaviness When Producing Contemporary Metal Music." *Journal on the Art of Record Production*. June 2012.
- Nobile, Drew. "Counterpoint in Rock Music." *Music Theory Spectrum*. 37: 2 (Fall 2015), pp. 189-203.
- _____. "Verse-Chorus Forms as Harmonic Patterns." Paper presented at the annual meeting of the Society for Music Theory, St. Louis, Missouri, October 29-November 1, 2015.
- Osborn, Brad. "Subverting the Verse-Chorus Paradigm: Terminally Climactic Forms in Recent Rock Music." *Music Theory Spectrum*. 35:1 (Spring 2013), pp. 23-47.
- _____. "Understanding Through-Composition in Post-Rock, Math-Metal, and other Post-Millennial Rock Genres." *Music Theory Online*. 17:3. October 2011.
- Seabrook, John. "The Doctor is In: A Technique for Producing No. 1 Songs." *The New Yorker*. October 2013. Web. 14 February 2016.
- Seay, Toby. "Capturing that Philadelphia Sound: Technical Exploration of Sigma Sound Studios." *Journal on the Art of Record Production*. June 2012.
- Senior, Mike. *Mixing Secrets for the Small Studio*, Oxford, Elsevier Inc., 2011.
- Shelvock, Matt. "Interview with Ben Fowler." *Journal on the Art of Record Production*. June 2012.
- Solberg, Ragnhild T. "Waiting for the Bass to Drop: Correlations Between Intense Emotional Experiences and Production Techniques in Build-up and Drop Sections of Electronic Dance Music." *Dancecult: Journal of Electronic Dance Music Culture* 6:1. December 2014, pp. 61-82.

- Spicer, Mark. "(Per)Form in(g) Rock: a Response." *Music Theory Online*. 17:3. October 2011.
- _____. Review of Walter Everett, *The Beatles as Musicians: The Quarry Men Through Rubber Soul*. *Music Theory Online*. 11:4. October 2005.
- Summach, Jay. "The Structure, Function, and Genesis of the Prechorus." *Music Theory Online* 17:3. October 2011.
- Temperley, David. "The Cadential IV in Rock," *Music Theory Online* 17:1. April 2011.
- Terasawa, H. *A Hybrid Model for Timbre Perception: Quantitative Representations of Sound Color and Density*. PhD Thesis, Stanford University. 2009.
- Tingen, Paul. "Dave Ogilvie Mixing 'Call Me Maybe'." *Soundonsound*. September 2012. Web. 13 May 2013.
- _____. "Fabian Marasciullo: Recording Lil Wayne's 'How to Love'." *Soundonsound*. December 2011. Web. 13 May 2013.
- _____. "François Tétaz: Mixing Gotye's 'Somebody That I Used To Know'." *Soundonsound*. July 2012. Web. 13 May 2013.
- _____. "Reverse Engineering." *Soundonsound*. October 2012. Web. 13 May 2013.
- _____. "Robopop: Producing Lana Del Rey's 'Videogames'." *Soundonsound*. April 2012. Web. 13 May 2013.
- _____. "Trevor Muzzy: Recording Nicki Minaj's 'Starships'." *Soundonsound*. August 2012. Web. 13 May 2013.
- Wessel, David. "Timbre Space as a Musical Control Structure." *Computer Music Journal*, 3:2 (June 1979), pp. 45-52.
- Williams, Alan. "Putting it on Display: The Impact of Visual Information on Control Room Dynamics." *Journal on the Art of Record Production*, June 2012.
- Zak, Albin. *The Poetics of Rock Composition: Multitrack Recordings as Compositional Practice*, Ph.D. Dissertation, The City University of New York, 1997.

RECORDINGS

- Aguilera, Christina. "Your Body." *Lotus*. RCA Records 88765404212. 2012.
- Alesso. "Heroes" featuring Tove Lo. *Forever*. Alefune/Def Jam Recordings B0023248-02. 2015.
- Becky G. "Break a Sweat." YouTube video. 3:33. Posted October 9, 2015.
<https://www.youtube.com/watch?v=yfuzpfpeX1s>.
- _____. "Can't Stop Dancing." YouTube video. 3:41. Posted December 3, 2014.
<https://www.youtube.com/watch?v=7NwzxGntDgs>.
- _____. "Lovin' so Hard." YouTube video. 4:11. May 6, 2015.
<https://www.youtube.com/watch?v=n6-RFMfPSX8>.
- _____. "Shower." YouTube video. 3:40. Posted June 30, 2014.
https://www.youtube.com/watch?v=50-_oTkmF5I.
- Beyoncé. "Single Ladies." *I am... Sasha Fierce*. Columbia 88697 19492 2. 2008.
- Bieber, Justin. "Beauty and a Beat" featuring Nicki Minaj. *Believe*. Def Jam Recordings B0016934-02. 2012.
- _____. "Boyfriend." *Believe*. Def Jam Recordings B0016934-02. 2012.
- _____. "Sorry." *Purpose*. Def Jam Recordings B0024042-02. 2015.
- _____. "What do You Mean." *Purpose*. Def Jam Recordings B0024042-02. 2015.
- Cara, Alessia. "Here." *Know-it-all*. Def Jam Recordings B0024071-02. 2015.
- Chainsmokers. "Roses" featuring Rosez. *Now That's what I Call Music! 58*. Sony 889853183029. 2016.
- Ciara. "Body Party." *Ciara*. Epic Records 88691971032. 2013.
- Cruz, Taio. "Dynamite." *Rokstarr*. Mercury B0014330-02. 2010.

Cyrus, Miley. "Adore You." *Bangerz*. RCA Records 88883-74523-2. 2013.

_____. "See You Again." *Breakout*. Hollywood Records D000212902. 2008.

_____. "We Can't Stop." *Bangerz*. RCA Records 88883-74523-2. 2013.

_____. "Wrecking Ball." *Bangerz*. RCA Records 88883-74523-2. 2013.

Daft Punk. "Get Lucky." *Random Access Memories*. Columbia 88883716862. 2013.

_____. "Lose Yourself to Dance." *Random Access Memories*. Columbia 88883716862. 2013.

Daya. "Hide Away." *Now That's what I Call Music! 57*. Sony 88875183312. 2015.

Del Rey, Lana. "Summertime Sadness." *Born to Die*. Interscope Records B0016425-02. 2012.

Derulo, Jason. "Want to Want Me." *Everything is 4*. Warner Bros. 549277-2. 2015.

_____. "Watcha Say." *Jason Derülo*. Warner Bros. 519657-2. 2010.

DJ Snake. "You Know You Like It" featuring AlunaGeorge. *Now That's what I Call Music! 55*. Universal Music Group 88875125062. 2014.

Drake. "Take Care" featuring Rihanna. *Take Care*. Cash Money Records B0016135-02. 2011.

Fifth Harmony. "Worth It" featuring Kid Ink. *Reflection*. SYCO 88843094742. 2015.

Goulding, Ellie. "Burn." *Halcyon Days*, reissue. Interscope Records B0019011-02. 2013.

_____. "Lights." *Lights*. Interscope B001532902. 2011.

Grande, Ariana. "Almost is Never Enough" featuring Nathan Sykes. *Yours Truly*. Republic Records B0018929-02. 2013.

_____. "Break Free." *My Everything*. Republic Records B0021395-02. 2014.

- _____. “Dangerous Woman.” *Dangerous Woman*. Republic Records 00602547908483. 2016.
- _____. “Focus.” *Dangerous Woman*. Republic Records 00602547908483. 2016.
- _____. “One Last Time.” *My Everything*. Republic Records B0021395-02. 2014.
- _____. “Problem” featuring Iggy Azalea. *My Everything*. Republic Records B0021395-02. 2014.
- _____. “Right There” featuring Big Sean. *Yours Truly*. Republic Records B0018929-02. 2013.
- _____. “The Way” featuring Mac Miller. *Yours Truly*. Republic Records B0018929-02. 2013.
- _____. “Love Me Harder” featuring The Weeknd. *My Everything*. Republic Records B0021395-02. 2014.
- Gomez, Selena. “Good for You.” *Revival*. Interscope Records B0024000-02. 2015.
- _____. “Same Old Love.” *Revival*. Interscope Records B0024000-02. 2015.
- Gotye. “Somebody that I Used to Know” featuring Kimbra. *Making Mirrors*. Sample 'N' Seconds Records B0016449-02. 2011.
- Guetta, David. “Hey Mama” featuring Nicki Minaj. *Listen*. Atlantic Records 546510-2. 2014.
- Guns N’ Roses. “Sweet Child O’ Mine.” *Appetite for Destruction*. Geffen Records 9 24148-2. 1987.
- Harris, Calvin. “Outside” featuring Ellie Goulding. *Motion*. Columbia 88875008972. 2014.
- Hozier. “Take Me to Church.” *Hozier*. Columbia 88843099962. 2013.
- Icona Pop. “All Night.” *This Is--*. Big Beat Records 536110. 2013.
- _____. “I Love It” featuring Charlie XCX. *This Is--*. Big Beat Records 536110. 2013.

Jack Ü. "Where Are Ü Now" featuring Justin Bieber. *Skrillex and Diplo Present Jack Ü*. Warner Music 7576866966. 2015.

Jepsen, Carly Rae. "Call Me Maybe." *Kiss*. Interscope Records B0017363-02. 2012.

Lady Gaga. "Bad Romance." *The Fame Monster*. Streamline Records B0013872-02. 2009.

Lamar, Kendrick. "Bitch, Don't Kill My Vibe." *good kid, m.A.A.d city*. Interscope Records B0018379-02IN02. 2012.

_____. "Swimming Pools." *good kid, m.A.A.d city*. Interscope Records B0018379-02IN02. 2012.

Lil Wayne. "She Will" featuring Drake. *The Carter IV*. Universal Republic B00015548-02. 2011.

Lo, Tove. "Habits." *Queen of the Clouds*. Island Records B0021921-02. 2014.

_____. "Not on Drugs." *Queen of the Clouds*. Island Records B0021921-02. 2014.

_____. "Talking Body." *Queen of the Clouds*. Island Records B0021921-02. 2014.

_____. "Timebomb." *Queen of the Clouds*. Island Records B0021921-02. 2014.

Lorde. "Royals." *Pure Heroine*. Republic Records B0019254-02. 2013.

Lovato, Demi. "Cool for the Summer." *Confident*. Hollywood Records D002293902. 2015.

Ludacris. "Representin" featuring Kelly Rowland. *Ludaversal*. Def Jam Recordings B002288102. 2015.

Major Lazer and DJ Snake. "Lean On." *Peace is the Mission*. Mad Decent MDCT26.2. 2015.

Michael, George. "Freedom '90." *Listen Without Prejudice Vol. 1*. Columbia Records CK 46898. 1990.

Miguel. "Adorn." *Kaleidoscope dream*. RCA Records 88725-47203-2. 2012.

_____. "How Many Drinks." *Kaleidoscope dream*. RCA Records 88725-47203-2. 2012.

- Mikael, Sebastian. "Last Night" featuring Wale. *Speechless*. Slip 'n' Slide Records 2537838752. 2014.
- Minaj, Nicki. "The Night is Still Young." *The Pinkprint*. Cash Money/Republic B0022248-02. 2015.
- _____. "Starships." *Pink Friday: Roman Reloaded*. Universal Motown B001512402. 2010.
- Perry, Katy. "E.T." *Teenage Dream*. Capitol Records 509996 84601 2 9. 2010.
- _____. "California Gurls." *Teenage Dream*. Capitol Records 509996 84601 2 9. 2010.
- _____. "I Kissed a Girl." *One of the Boys*. Capitol Records 509995 04249 2 4. 2008.
- _____. "Roar." *Prism*. Capitol Records B001921602. 2013.
- _____. "Teenage Dream." *Teenage Dream*. Capitol Records 509996 84601 2 9. 2010.
- _____. "The One that Got Away." *Teenage Dream*. Capitol Records 509996 84601 2 9. 2010.
- _____. "Unconditionally." *Prism*. Capitol Records B001921602. 2013.
- Pitbull and Ne-Yo. "Time of Our Lives." *Globalization*. RCA Records 88875-01119-2. 2014.
- Queen. "Crazy Little Thing Called Love." *Greatest Hits*. Hollywood Records HR-61265-2. 1974.
- Rihanna. "American Oxygen." YouTube video. 5:20. Posted April 16, 2015.
<https://www.youtube.com/watch?v=Ao8cGLIMtvg>.
- _____. "Pour it Up." *Unapologetic*. Def Jam Recordings B0017811-02. 2012.
- _____. "Stay" featuring Mikky Ekko. *Unapologetic*. Def Jam Recordings B0017811-02. 2012.
- _____. "We Found Love" featuring Calvin Harris. *Talk That Talk*. Def Jam Recordings. B0016313-02. 2011.

- _____. "Work" featuring Drake. *Anti*. Roc Nation B002299302. 2016.
- Rowland, Kelly. "Motivation" featuring Lil Wayne. *Here I Am*. Universal Motown B0014495-02. 2010.
- Sia. "Chandelier." *1000 Forms of Fear*. Monkey Puzzle Records 88843-07404-2. 2014.
- _____. "Elastic Heart." *1000 Forms of Fear*. Monkey Puzzle Records 88843-07404-2. 2014.
- Spears, Britney. "3." *The Singles Collection*. Jive 88697-59675-2. 2009.
- _____. *Blackout*. Jive 8869 719073-2. 2007.
- _____. "Hold it Against Me." *Femme Fatale*. Jive 88697-85332-2. 2011.
- _____. "I Wanna Go." *Femme Fatale*. Jive 88697-85332-2. 2011.
- _____. "Till the World Ends." *Femme Fatale*. Jive 88697-85332-2. 2011.
- _____. "Toxic." *In the Zone*. Jive 82876-53748-2. 2003.
- Summer, Donna. "I Feel Love." *I Remember Yesterday*. Casablanca NBLP 7056. 1977.
- Swift, Taylor. "Blank Space." *1989*. Big Machine Records BMRBD00500A. 2015.
- _____. "I Knew You Were Trouble." *Red*. Big Machine Records BMR310400A. 2012.
- _____. "Out of the Woods." *1989*. Big Machine Records BMRBD00500A. 2015.
- _____. "Shake it Off." *1989*. Big Machine Records BMRBD00500A. 2015.
- _____. "Style." *1989*. Big Machine Records BMRBD00500A. 2015.
- _____. "We are Never Ever Getting Back Together." *Red*. Big Machine Records BMR310400A. 2012.
- _____. "Wildest Dreams." *1989*. Big Machine Records BMRBD00500A. 2015.
- _____. "You Belong with Me." *Fearless*. Big Machine Records. BMRATS0200. 2008.

Thicke, Robin. "Back Together" featuring Nicki Minaj. YouTube video. 3:48. Posted August 13, 2015. <https://www.youtube.com/watch?v=xqtUuHFAtzg>.

Trainor, Meghan. "All About that Bass." *Title*. Epic 88875020882. 2014.

Usher. "DJ Got Us Fallin' in Love" featuring Pitbull. *Raymond vs. Raymond*. LaFace Records 88697-61552-2. 2010.

_____. "More." *Raymond vs. Raymond*. LaFace Records 88697-61552-2. 2010.

Weeknd, The. "Earned It." *Fifty Shades of Gray: Original Motion Picture Soundtrack*. Republic Records B0022595-02. 2015.

_____. "Wicked Games." *House of Balloons*. Republic Records B0023686-02. 2011.

Will.I.Am. "Scream and Shout" featuring Britney Spears. *#Willpower*. Interscope Records B0018251-02. 2013.

Zayn. "Pillowtalk." *Mind of Mine*. RCA Records 88985307902. 2016.